Università di Genova
DIPARTIMENTO INFORMATICA, BIOINGEGNERIA, ROBOTICA E INGEGNERIA DEI SISTEMI

Computer Science Workshop
PhD program in Computer Science and Systems Engineering

Efficient Unsupervised Learning for Plankton Images

Paolo D. Alfano, Marco Rando, Marco Letizia, Lorenzo Rosasco, Francesca Odone, Vito P. Pastore

Introduction

Monitoring plankton populations in situ is fundamental to preserve the aquatic ecosystem. Plankton microorganisms are susceptible of minor environmental perturbations[1] and can be regarded as biosensors. In this context, the adoption of machine learning algorithms may be affected by the significant cost of manual annotation. To address these challenges, we propose an efficient unsupervised learning pipeline for classification of plankton microorganisms. A Variational Autoencoder (VAE)[2] is trained on features extracted by a pre-trained neural network. We then use the learnt latent space as input for clustering. We compare our method with state-of-the-art unsupervised approaches, where a set of pre-defined hand-crafted features is used for clustering of plankton images.

Methodology

A Variational Auto Encoder (VAE) learns to reproduce input while learning a lower dimensional representation by encoding the input into a multivariate latent distribution:

$$x' = d(e(x))$$

$$l(x, x') = \|x - x'\|^2 + D_{KL}(\mathcal{G}(\mu(x), \sigma(x)), \mathcal{N}(0, I)).$$

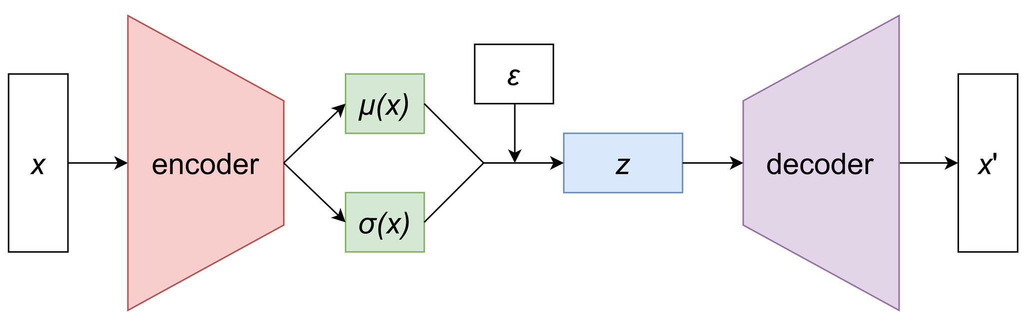

Figure 1: Schematic VAE representation

Our pipeline includes the following steps:

1. Pre-processing: images are resized and normalized.
2. Features extraction: images are given as input to a deep neural network pre-trained on the ImageNet dataset, without further tuning.
3. Dimensionality reduction and clustering: train a convolutional Variational Auto Encoder. Learnt latent space is fed to a clustering algorithm.

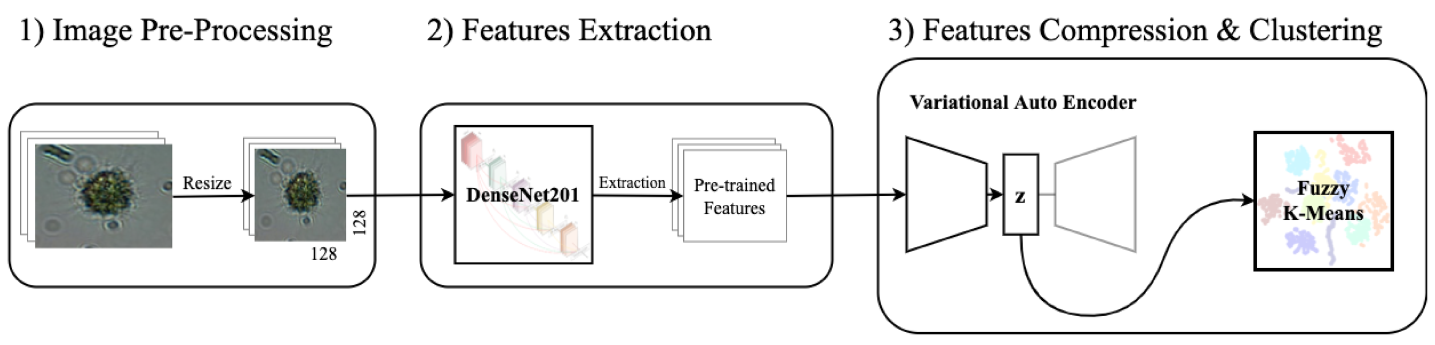

Figure 2: The proposed pipeline

Experiments

Datasets:

- Lensless: 10 classes, with 640 color images each.
- WHOI40: 40 classes, each represented by 100 grayscale images.
- WHOI22: 22 fine grained species, 300 grayscale images each.

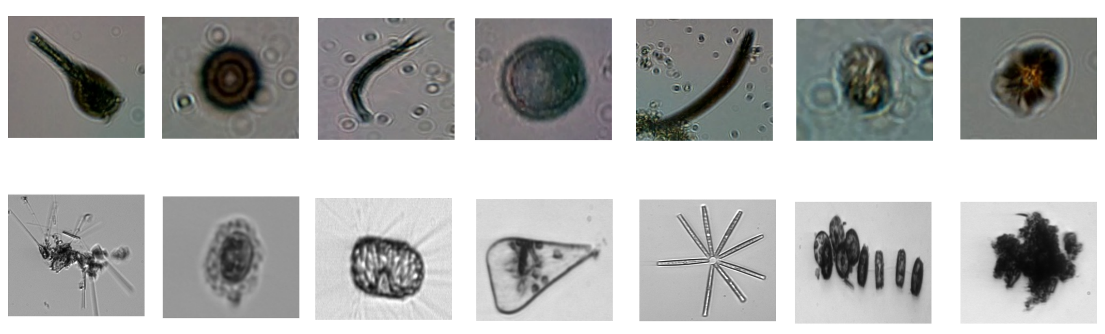

Figure 3: Sample images. Lensless (top), WHOI (bottom)

We evaluate our pipeline with the *purity* measure:

$$\text{purity}(\Omega, C) = \frac{1}{N} \sum_k \max_j |w_k \cap c_j|,$$

where $\Omega = \{w_1, ..., w_k\}$ set of clusters, $C = \{c_1, ..., c_j\}$ set of ground-truth classes. We can show qualitative benefit using pre-trained features on the lensless dataset.

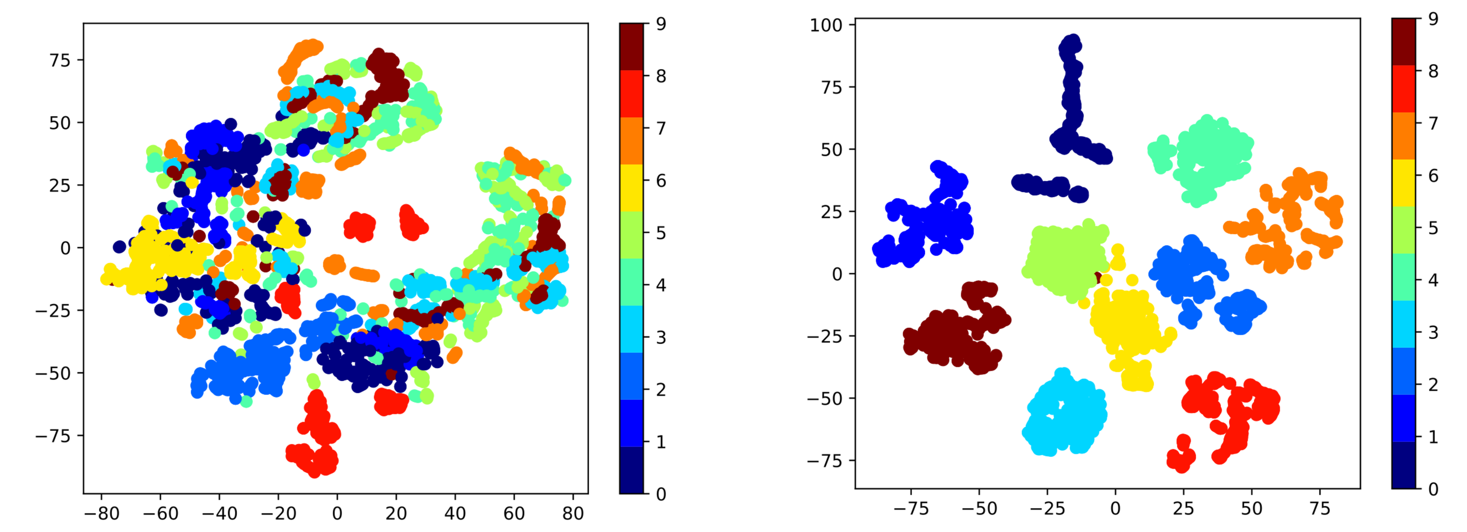

Figure 4: Learnt latent space via TSNE with different input. Images (left). Pre-trained features (right)

Quantitative benefit of using pre-trained features on the lensless dataset:

Algorithm/Z	10	30	50	100	500
image-VAE	0.53 ± 0.017 (1.4 ± 0.5)	0.55 ± 0.04 (1.6 ± 0.49)	0.58 ± 0.01 (2.0 ± 0.63)	0.59 ± 0.01 (1.6 ± 0.48)	0.62 ± 0.01 (2.0 ± 0.0)
FE-VAE	0.98 ± 0.01 (0.0 ± 0.0)	0.98 ± 0.03 (0.0 ± 0.0)	0.98 ± 0.01 (0.0 ± 0.0)	0.98 ± 0.02 (0.0 ± 0.0)	0.98 ± 0.02 (0.0 ± 0.0)

- Pre-trained features increases the purity with an average improvement of 30%
- Using a bigger latent space does not imply a performance improvement

Quantitative results on the other datasets:

Dataset/Z	10	30	50	100	500
WHOI 40	0.66 ± 0.01 (5.8 ± 0.75)	0.71 ± 0.02 (5.8 ± 1.16)	0.73 ± 0.02 (5.2 ± 0.98)	0.77 ± 0.01 (3.8 ± 1.16)	0.77 ± 0.01 (4.0 ± 0.63)
WHOI 22	0.63 ± 0.004 (2.0 ± 0.63)	0.66 ± 0.01 (1.6 ± 0.49)	0.68 ± 0.005 (1.6 ± 0.49)	0.68 ± 0.006 (1.4 ± 0.5)	0.68 ± 0.01 (1.8 ± 0.4)

- Best performances correspond to a latent space size $Z = 100$.

We benchmarked our results using a state-of-the-art unsupervised learning pipeline.

Algorithm/Dataset	Lensless	WHOI 40	WHOI 22
Pipeline from [3]	0.93 (0) [3]	0.71 (5) [3]	0.56 (3)
Ours	**0.98 (0)**	**0.77 (4)**	**0.68 (2)**

Conclusions

We introduced an efficient unsupervised pipeline for plankton images. Input images are fed to a pre-trained neural network. Output features are used as inputs to train a Variational Auto Encoder. The latent space representation is used by clustering algorithm.
A VAE with a latent space dimension 100 with pre-trained input features gives the best results for all the datasets included in our work. By using a set of pre-trained features, our pipeline represents an efficient approach in terms of time and resources.
Future developments will extend our analysis to other datasets to test our pipeline on a more general context.

References

[1] Daniel Boyce, Marlon Lewis, and Boris Worm. Global phytoplankton decline over the past century. *Nature*, 466:591–6, 07 2010.

[2] Diederik P Kingma and Max Welling. Auto-encoding variational bayes, 2014.

[3] Vito P. Pastore, Thomas G. Zimmerman, Sujoy K. Biswas, and Simone Bianco. Annotation-free learning of plankton for classification and anomaly detection. *Scientific Reports*, 10(1):12142, 2020.

CONTACTS

Paolo D. Alfano
paolodidier.alfano@
edu.unige.it

Marco Letizia
marco.letizia@edu.unige.it

Marco Rando
marco.rando@edu.unige.it

Vito P. Pastore
vito.paolo.pastore@unige.it

CSW