

Pre-processing proposals for plankton images classification

Lorenzo Serafini¹,
Michele Sprocatti², Alberto Pasqualetto³

Department of Information Engineering, University of Padua
Via Gradenigo 6, 35131 Padova, Italy

{lorenzo.serafini.1¹, michele.sprocatti², alberto.pasqualetto.2³}@studenti.unipd.it

Abstract

This paper introduces a new image pre-processing technique specifically designed for classifying plankton images with AlexNet. It enforces the use of the unsharp method and the application of padding before resizing the images in input to the network. It compares the results with other pre-processing methods found in the literature.

Introduction

Due to their tiny size, plankton often appears in low-quality and low-resolution images. This makes automatic classification challenging. Thus pre-processing techniques can be employed to overcome these limitations by highlighting the most important features of the plankton, such as shape and texture. In the end, it will result in a better-performing input for the neural network.

Dataset

ZooScan (Gorsky et al. 2010) comprises 3771 images captured in the Bay of Villefranche-sur-mer using Zooscan technology. These images are classified into 20 categories, with varying sample sizes ranging from 28 to 427 images per category. The majority of categories encompass zooplankton, various Medusae species, and zooplankton eggs, while the remaining categories consist of non-zooplankton organisms and images with poor focus.

The dataset is split into two folds: for each fold the net is trained with 80% of the patterns and the remaining 20% is used to evaluate the model.

The images have different dimensions but all of them consist of pixels with an 8-bit depth and three color channels. The background of all images is consistently defined by the white color, expressed as RGB(255, 255, 255).

Network

We fed the images into an AlexNet (scheme in figure 4) where we replaced the last 3 layers with a fully connected layer using a SoftMax activation function. The network requires input images with dimension $227 \times 227 \times 3$. It was trained on the provided dataset for 30 epochs.

Method

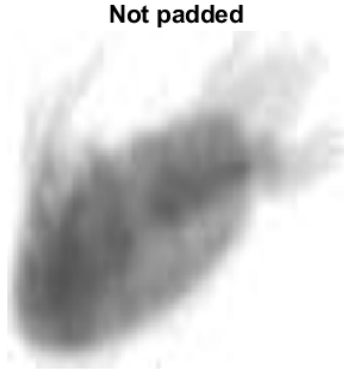
When we apply the resize operation, since some images have a dimension significantly larger than others, we squeeze the image and lose the plankton morphology: this can lead to a misrepresentation of the pattern. A comparison of an image resized with and without padding is shown in figure 1. To avoid incurring this problem, we added white padding to the smaller dimension of the image obtaining a square image of $\max(\text{height}, \text{width}) \times \max(\text{height}, \text{width})$ size. In this way, the resizing to 227×227 would not cause the squeezing effect.

Our idea for the preprocessing of the plankton images is to increase the contrast of the image in order to better visualize the details and enhance the transitions in it. Applying histogram equalization to the whole image was not effective since a major part of the image consists of the white background: for this reason the gray and darker pixels were set to black or very close to it after the equalization, obtaining the opposite effect instead of enhancing the differences in pixel values. A comparison of an image with and without histogram equalization is shown in figure 2. Knowing this behavior we tried to apply local methods like Local Histogram Equalization and local laplacian filtering, but both of them did not improve the accuracy of the network, on the contrary, the results were worse than the ones obtained without any pre-processing (accuracy of about 0.82): for LHE we achieved an accuracy of about 0.78, for local laplacian filtering an accuracy of about 0.79.

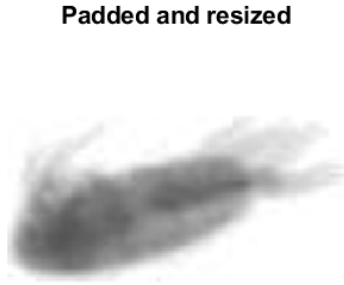
The Unsharp method is the one that worked best, the name comes from the fact that it produces a blurred version of the image and uses it to increase the contrast between neighboring pixels and highlight the transitions. This method blurs the image with a Gaussian filter, calculates the difference image between the original and the blurred one, and ultimately sums the original image with the difference image multiplied by a constant called "amount" following the formula:

$$\text{sharpened} = \text{original} + (\text{original} - \text{blurred}) \times \text{Amount}.$$

The Matlab `imsharpen` function's parameters *Radius* (the standard deviation of the Gaussian blur) and *Amount* have been set after a validation process where the training set is composed by 60% of the images and the validation set by 20% of the images, the remaining is used afterward to



(a) An image from the dataset resized to 227×227



(b) Same image from the dataset padded before resizing to 227×227

Figure 1: Comparison of an image resize with and without padding

evaluate the model. After the validation process, the dataset was split 80/20 as explained in section Dataset by merging the training and validation splits, finally, the net was trained with the best parameters found.

Results

As can be seen in table 1, Dai et al.'s pre-processing method is not effective using AlexNet architecture.

The Unsharp method, while not providing a significant improvement in accuracy, is the best pre-processing method among the ones tested.

On the other hand, applying padding before resizing the images has a positive effect on the accuracy of the net, increasing it by about 2%.



(a) An image from the dataset



(b) Same image from the dataset with histogram equalization applied: contrast is lost

Figure 2: Comparison of an image with and without histogram equalization

Method		Accuracy	
		No Padding	Padding
No pre-processing	Fold 1	0.81457	0.85298
	Fold 2	0.83311	0.83179
Dai et al. 2017	Fold 1	0.80397	0.85298
	Fold 2	0.81325	0.78808
Unsharp mask	Fold 1	0.81854	0.86490
	Fold 2	0.83179	0.83576

Table 1: Accuracy of the net with various pre-processing methods

The resulting confusion matrices are shown in figure 3.

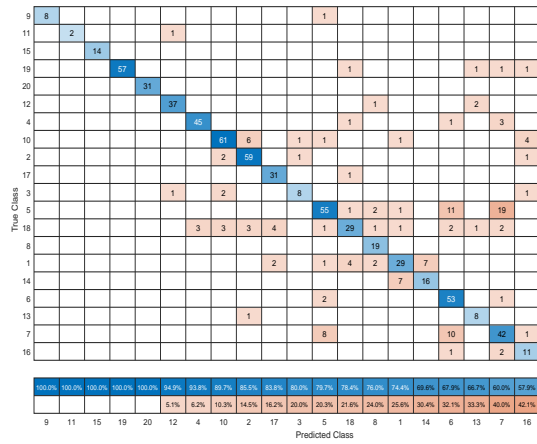
Conclusions

In this paper, we presented a new image pre-processing approach for classifying plankton images with AlexNet. The results show that Unsharp is the best pre-processing method that we tested for increasing the accuracy of the network even though not providing a very significant gain in performance. The usage of padding, instead, is the key to gaining some percentage points in accuracy.

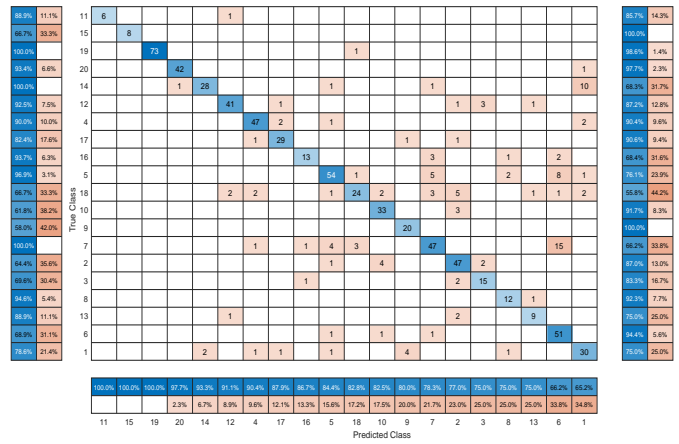
The combined application of padding and unsharp method allows the network to reach an average accuracy of 85% on both folds of the given dataset.

References

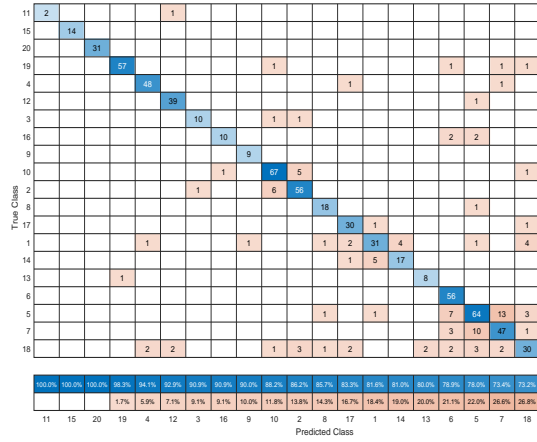
- Dai, J.; Yu, Z.; Zheng, H.; Zheng, B.; and Wang, N. 2017. A Hybrid Convolutional Neural Network for Plankton Classification. In Chen, C.-S.; Lu, J.; and Ma, K.-K., eds., *Computer Vision – ACCV 2016 Workshops*, 102–114. Cham: Springer International Publishing. ISBN 978-3-319-54526-4.
- Gorsky, G.; Ohman, M. D.; Picheral, M.; Gasparini, S.; Stemmann, L.; Romagnan, J.-B.; Cawood, A.; Pesant, S.; García-Comas, C.; and Prejger, F. 2010. Digital zooplankton image analysis using the ZooScan integrated system. *Journal of Plankton Research*, 32(3): 285–303.



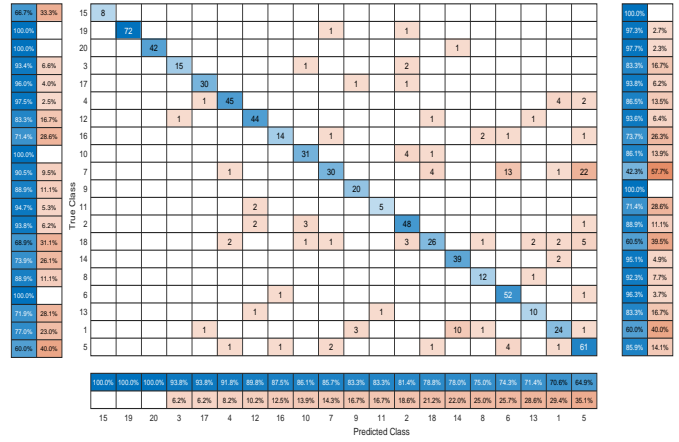
(a) Confusion Matrix fold 1 without pre-processing



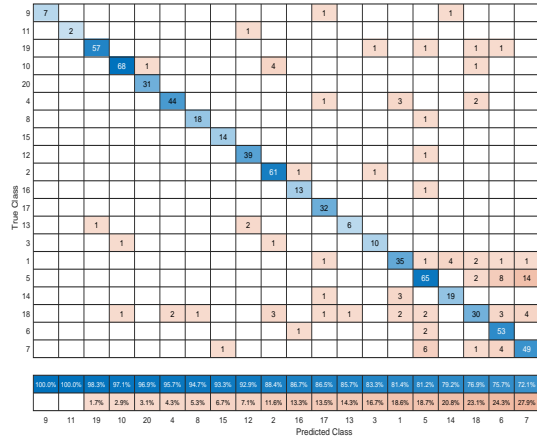
(b) Confusion Matrix fold 2 without pre-processing



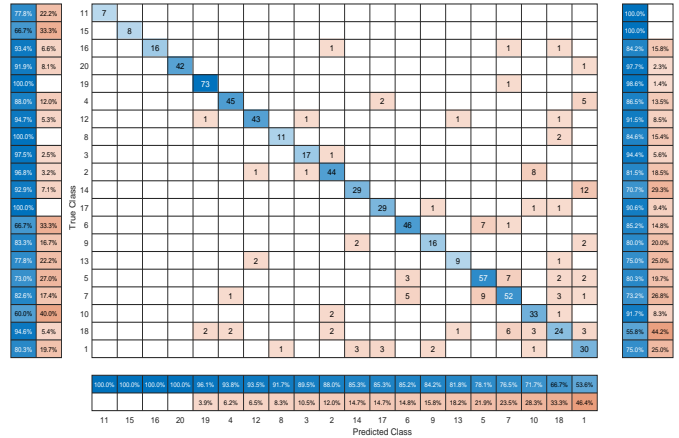
(c) Confusion Matrix fold 1 with square padding



(d) Confusion Matrix fold 2 with square padding



(e) Confusion Matrix fold 1 with Unsharp mask and padding



(f) Confusion Matrix fold 2 with Unsharp mask and padding

Figure 3: Confusion matrices

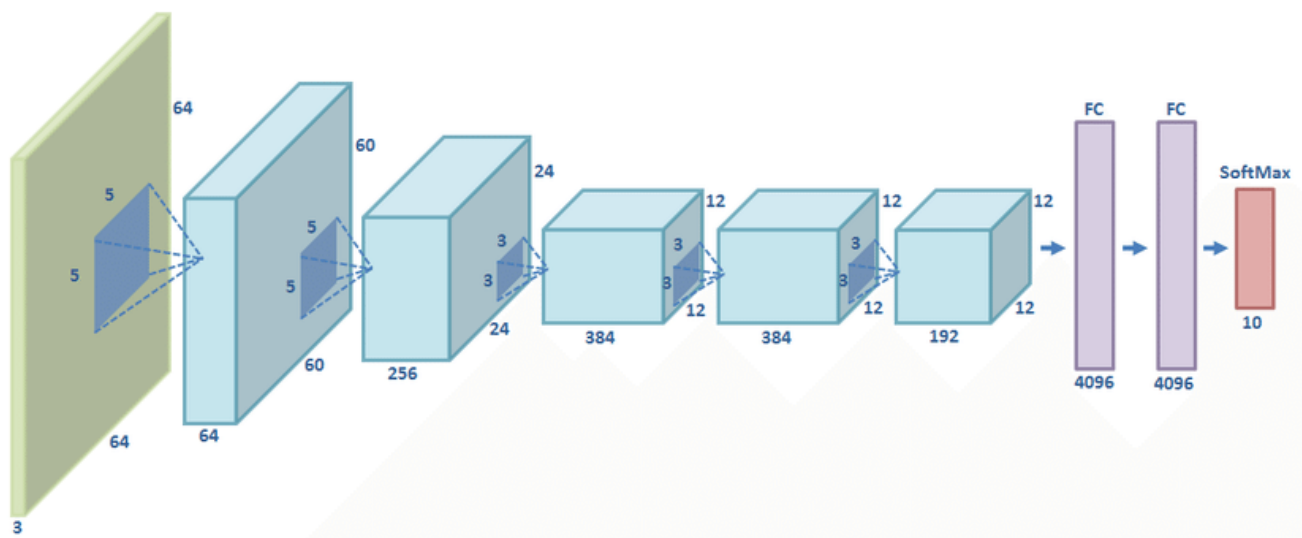


Figure 4: AlexNet