Image Classification of Aquatic Macroinvertebrates

ALEX JUSTESEN KARLSEN
Aarhus University - Department of Engineering
au515295@post.au.dk

April 26, 2019

Abstract

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

1. INTRODUCTION

Life below water is a critical ecosystem, that experiences a lot of challenging changes, which is way it has found its way to the UN Sustainable Goals [1]. The Finnish Environment Institute [2] SYKE claims, that the loss of aquatic biodiversity and associated Ecosystem Services is one of the most pressing problems on Earth. The DETECT project aims to identify the various species of aquatic macroinvertebrates by imagery and classification using computer vision and machine learning.

This project is a part of the Computer Vision and Machine Learning course at Aarhus University. The project is run as a Kaggle Challenge with the aim of improving the classification accuracy on the data set. The data set is available for download on the challenge page on Kaggle. The data set consists of raw images of 29 different classes of aquaitc macroinvertebrates and is split in training-, validation- and test set. Training- and validation set consist of correspendingly 5830 and 2298 labeled images and the test set consists of 3460 unlabeled images. The results of the test set are uploaded to Kaggle, where the results are scored by comparing with withhold ground truth labels.

All the data set are available as feature vectors, that have been extracted from a pretrained AlexNet. AlexNet [3] was the first convolutional neural network (CNN) used on the ImageNet challenge, and since then CNNs have become the golden standard to obtain remarkable classification results. The AlexNet features for this project are used to train both linear and non-linear classifiers to evaluate different classification methods on this data set and get hands-on experience with a real classification challenge.

Additionally the project examines a more modern CNN, InceptionV3 [4] pretrained on the ImageNet data set, is used as a feature extractor. The same classification methods are used to evaluate importance of features extraction og images.

Among the different classification techniques used are; K-Nearest Neighbor (kNN), Linear Discriminat Analysis (LDA), Support Vector Machines (SVM) and Neural Networks (NN). Principal Component Analysis (PCA) are used for dimensionalty reduction and Search Grid are used for hyper-parameter optimization.

The rest of the paper is structured as follows; section 2 desribes the data set, section 3 describes the methods and frameworks used in the project, section 4 describes the result of the

project, section 5 is a discussion of the results and last in section 6 the project is concluded upon.

2. Data Set

The FIN-Benthic [1] dataset 2 are used in this paper. The dataset consists of 29 different classes of aquatic macroinvertebrates. The dataset counts 11,588 images divided into training set (5830 images), validation set (2298) and test set (3460 images).

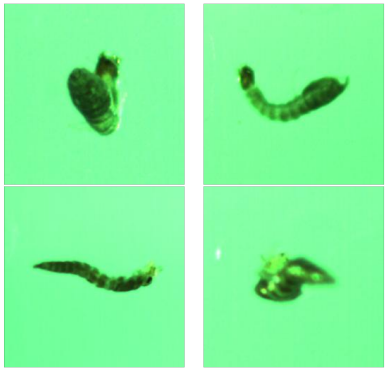

Figure 2.1. *Example of aquatic macroinvertebrates*

The training and validation set are annotated with class labels, whereas the test set are not. This paper is a part of a Kaggle [link ref] Challenge with the aim of getting the best classification score on the test set, that are evaluated on Kaggle. For the challenge feature vectors produced by an AlexNet[ref] training on the raw images. The features are the output of the last convolutional layer and flattened into a one-dimnesional vector with 4096 features. All images and vectors comes in pairs of images of the same bug taken form perpendicular angles, an attempt to take advantage of the systematic series the images are all images vertically stacked. If we have a pair of x_1 and x_2, then x_1 and x_2 are stacked as a combination of the two;

$$x_1 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, x_2 = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$$

[1] https://etsin.avointiede.fi/dataset/urn-nbn-fi-csc-kata20170615175247247938

The vectors now contains $2 \times 4096 = 8192$ features. The concatenated version of the dataset are referred to as FIN-Benthic-Concatenated. Futhermore the training and validation sets are merged to form a combined set, to enable Cross-Validation for the different classification algorithms tested in this paper. Cross-Validation and Machine Learning methods used are described in the next section.

3. Methods

Several experiments on this supervised learning challenge have been conducted using a variety of different techniques for classification are beformed on the two encoded datasets; FIN-Benthic and FIN-Benthic-concatenated. The classification methods are; k-Nearest Neighbors (kNN), Linear Disriminant Analysis (LDA), Support Vector Machines (SVM), Perceptron and Neural Networks (NN). Cross Validation (CV) are use for model selection to evaluate the best performing model. All the methods have been applied to the datasets on full dimensions, additionally they have been applied with dimensionality reduction as a preprocessing step in the classification pipeline. Grid search have been used to search for good hyperparameters for all classifiers and dimensionality reduction techniques. Lastly existing Convolutional Neural Network architectures have been used on the raw images and using the promise of transfer learning to use a pretrained model to gain faster and better performance on a small dataset. For the transfer learning task the existing train, validation and test split have been used.

k-Nearest Neighbors (kNN) K-Nearest Neighbor is a distance-based classifier and as the name implies classifies to the most common occuring class of the k-nearest neighbors. There exist a variety of distance metrics, some example are; euclidean, minikowski, manhanobli and manhattan distance.

Linear Discrmininat Analysis (LDA) Linear Discrmininat Analysis is a linear transformation

of the data. The linear transformation is an optimization, that tries to maximize the distance between classes and minimize the distance within a class. The classifier can also be used for dimensionality reduction, as one must specify the n number of components i.e. dimensions to transform the data onto, if the number of dimensions are less than the original data, than the dimensionality is reduced.

elaborate

Support Vector Machine (SVM)

Perceptron

Neural Network (NN)

Convolutional Neural Network (CNN)

Dimensionality Reduction

Principal Component Analysis (PCA)

Model Selection Model selection are procedures to train the best model and avoid overfitting the model to the training data. Overfitting is the pitfall of having a model, that desribes the training data too well, thus suffers to not generalising the true underlying structure of the data. Although other methods exists the two used in the paper are Validation Holdout and Cross Validation.

Validation Holdout Validation Holdout takes a fraction of your available labelled data and hold it out of the training procedure. The validation subset is then used to validate the model's performance on data it has not been training. The assumption is, that the holdout set resembles the true data, and if the model performs well on both training and validation it is assumed to be a good model. The split is typically made 7030 for training data and validation data respectively.

Cross Validation (CV) Cross Validation extends on the holdout validation. The entire data set are split in k folds, for every k step one subset is for validation and the remaining $k-1$ fold are used for training. Figure illustrates the idea.

The ensures that the model have been trained on all the data and validated as well. The selection of k are typically either 5 or 10.

Hyperparameter Search There are mainly two approaches to search for hyperparameters; Random Search and Grid Search other than heuristic guessing. In this paper Grid Search are used.

Grid Search Grid Search is an approach to seek better hyperparameters for your model. You specify a set of options for all hyperparameters and then seacrh all possible combination in the grid. The down-side of this approch is, that it is still based on intuition of good hyperparameters, however it allows you to explore a broader space of setting at the cost of additional time. Nonetheless one combination, that was not considered might turn out to be the best one. The random approach are also widely used and might be preffered, as it may find a sweet spot, that goes beyond human intuition, however that is no guarantee and it is still time costly.

4. Results

5. Discussion

6. Conclusion

7. REFERENCES

[1] "Transforming our world : the 2030 agenda for sustainable development," *UN General Assembly*, accessed: 8 April 2019.

[2] K. Meissner, "Detect." [Online]. Available: https://www.syke.fi/en-US/Research__Development/Research_and_development_projects/Projects/DETECT

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, ser. NIPS'12. USA: Curran Associates Inc., 2012, pp. 1097–1105. [Online]. Available: http://dl.acm.org/citation.cfm?id=2999134.2999257

[4] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," 06 2016.