MICROBLINK

See beyond.

Mozgalo: Robust ML Challenge (ML - machine learning)

The description of the prize assignment

- **What exactly is the prize assignment**

In the last couple of years, machine learning has in a way revolutionized the field of computer vision. On some of the publicly available datasets, results surpassing human abilities have been achieved which indicates a high potential for serious industrial applications. On the other hand, the development of concrete solutions in practice is often significantly more complex due to additional requirements and limitations. The aim of this year's Mozgalo competition is to introduce students with the challenges of developing robust machine learning solutions. In theory, the assignment is a simple problem of image classification which becomes highly complex in practice. More specifically, the task is image classification of shopping mall receipts based on their visible logo characteristics. The teams shall develop the solution on a labelled set of 45 000 images and 25 different classes and their robustness will be measured on the set of images which simulate the real conditions of industrial applications.

- **Educational purpose**

After the initial success of deep learning, an expected point of saturation has been reached. As such, we have started to turn towards the questions of the robustness of learnt models, and with that, their interpretability. Currently, those are perhaps the greatest challenges of computer science where, on one end, the academia is trying to solve the issue of their interpretability and, on the other end, the industry is trying to solve the issues of their robustness. In both cases, the first step is to understand the limitations of current solutions. The aim of this year's Mozgalo is to motivate the students to make that first step in order to develop a much needed knowledge basis for high-quality, innovative work in the field of machine learning be it in academia or industry. The competitors will be introduced to the problems of robust optimization and have an opportunity to firsthand learn all that

is possible with the currently most advanced solutions of machine learning, but also the shortcomings of those solutions.

- **Potential applications**

The applications of the knowledge and experience acquired in this competition is very broad whether we are talking about academic research or development of concrete solutions in industry. The final solution for the given assignment should be able to identify the shopping mall the receipt belongs to based on the image of the receipt. It is also the first step of a system for automated data extraction from an image of a buyer's receipt. The cheap and quick data collection directly from buyers opens the doors to the highly advanced methods of market research without having any knowledge of the business of any single shopping mall.

- **The potential problems one may encounter while solving the prize assignment**

The assignment is demanding; it is necessary to know machine learning and computer vision well. The aim of workshops is to clear as many obscurities and to direct the teams in the right direction. As in previous competitions, there is a problem of computer resources. Ideally, the teams should have access to servers for learning in order to quickly carry out their experiments. Unfortunately, it has not been possible to organize it this year as well. However, the assignment was conceived with the emphasis on developing a smart network architecture which would be trained on a small set of date and which should give better results than a naive network architecture trained on a large amount of data. With that, we believe we have decreased the need for a large amount of computer resources.

Recommended programming languages and/or tools for solving the prize assignment

TensorFlow and PyTorch tools for machine learning.

An example of the dataset which contestants will receive

Training dataset:

- Altogether 45 000 images of receipts
- 25 different classes of shopping malls
- The total size of the file is 10GB.

Test dataset:

- Altogether 20 0000 images of different type.
- The set contains:
  - Images of receipts belonging to the same classes as those from the training dataset on which different transformations simulating real world use.
  - Images of random content collected through method of hard example mining with existing Microblink solutions.
- The total size of the file is 5 GB.

The description of input variables competitors will get

The competitors will get images with related classes given as string values.

Does the Client have experience in solving the same or similar assignment?

Microblink specializes in development of precise and quick solutions of machine learning and puts emphasis precisely on robustness. For this year's Mozgalo competition, the preparations began after the last year's competition with the aim of finding the topic which would be most useful for the students. The topic we choose was one which was minutely researched by Microblink Research and for which, one possible solution is currently used for receipt classification in the US market. In the meantime, we have participated in two international conferences, NIPS (Neural Information Processing Systems) and ICDAR (International Conference on Document Analysis and Recognition), where we had the opportunity to further elaborate on a topic and confirm its significance for the students. Concretely, we partook in lectures, workshops, and panel discussions with the topic of robust optimization and the shortcomings of the current solutions of machine learning.

Additional information

Understanding Adversarial Training: Increasing Local Stability of Neural Nets through Robust Optimization

https://arxiv.org/abs/1511.05432

Semi-Supervised Learning with Generative Adversarial Networks

https://arxiv.org/abs/1606.01583

Rethinking the Inception Architecture for Computer Vision

https://arxiv.org/abs/1512.00567

Spatial Transformer Networks

https://arxiv.org/abs/1506.02025

Attention in Neural Networks and How to Use It

http://akosiorek.github.io/ml/2017/10/14/visual-attention.html

Mandatory solution formats

- source code (Python, Java, C++)
- descriptive presentation of the solution (PPT, PDF)

Scoring distribution

The emphasis is on performance of the final solution, objectively scored using the standard macro f1 measure on the test set.

Criteria	**Points**	**Points share**
Performance	0-90	90 %
Ingenuity and applicability	0-10	10 %
Total	100	100%

At the beginning of the competition, each team will be given the train set and the first half of the test set they will use to develop their solution. Three times a day they can test their solution

accuracy by uploading the solution on the web service of the competition.

On the last day of the competition, teams will be given the other half of the test set for which they will have to upload their results in under 3 hours of receiving data. Obtained accuracy of the solution will not be publicly shown until the end of the competition, but it will determine the best N teams which made it to the final phase (the exact number of teams chosen for the final phase will be determined by the jury based on the differences between solutions).

Finalists will be given points based on the simple formula:

acquired points = 90 * test_set_macro_average_f1

Next there will be a discrete presentation of the solution to the jury where they will test the solution and check its legitimacy to confirm their finalists status.

In the final phase of the competition finalists will give a public presentation where they will be graded by the jury for ingenuity and applicability of their solution. Given points (max. 10) will be added to the previously acquired points which will then form the final sum. The winning team is the one with the highest number of points, followed by the second and the third team on the ranking list.

Additional benefits

Possibility of an internship at the Microblink research team.