Weakly-Supervised Learning
for Text Recognition

Supervisors:

Marçal Rossinyol   marcal@allread.ai
Dimosthenis Karatzas   dimos@cvc.uab.es

Company:

AllRead   http://www.allread.ai

Institute / Group:

Computer Vision Center   http://www.cvc.uab.cat/
Intelligent Reading Systems

Conditions:

Paid project depending on student experience. Possibility to do extra practicals at AllRead.

Keywords:

Scene Text Recognition, Incremental Learning, Weakly-Supervised Learning

ABSTRACT

This project deals with the design and implementation of weakly-supervised learning techniques for the specific case of scene text recognition.

Starting from a baseline model able to detect and recognize specific texts (such as vehicle license plates, container codes, wagon numbers, etc.), we want to apply weakly-supervised learning techniques to leverage inference data without requiring complex data annotations to incrementally train and enhance the models' performance.

Fig 1: *Sample of AllRead use-case scenarios.*

CONTEXT AND NEED

Text is omnipresent in man-made environments, and conveys important, semantic information about the world around us. When it comes to structured text, machine-based automated reading software can be the solution, but instead of a generic OCR a different approach is needed, adaptable and specialised to each reading domain. Such recognition models should be carefully tailored to work in specific scenarios (cf. Fig 1), guaranteeing an important recognition accuracy, but also, should be able to focus on the text that is deemed relevant for the user, ignoring the rest of non-relevant textual information appearing in the image.

However, in order to properly train such models, it takes tens or hundreds of thousands of training samples to optimize a deep neural network so that it is able to read accurately. Our objective is to build novel training procedures that drastically reduce such training data needs by several orders of magnitude.

In that aspect, having a weakly-supervised approach that harvests data at inference time would be of great value to leverage automatically annotated data in order to incrementally train baseline scene text recognition models.

PROJECT OBJECTIVES

The main objective of this research project is to design and implement a weakly-supervised learning approach for our scene text recognition baseline models.

Starting from those baseline models, the objective is to incrementally retrain them by using data with weak annotations obtained by processing thousands of videos coming from real operational environments, hoping that the models' performance will steadily increase as the models keep processing new data.

During this project, the student will be expected to:
- Study the state of the art and research scene text recognition models and weakly supervised techniques
- Take ownership of AllRead's baseline models and AllRead data
- Develop a novel weakly-supervised incremental learning technique
- Train, test and keep evolving the model while analysing accuracy
- Optimise the models for efficiency
- Evaluate the proposed methods with data from real and relevant environments

The project will be developed in Python using the Tensorflow 2.x framework.

AllRead will offer real-life test data and use-case scenarios on which the student will be able to establish pilot studies, to evaluate the efficiency of the produced models.

BACKGROUND AND TEAM

AllRead was founded in mid 2019 as a spinoff of the CVC to commercialise a novel text recognition technology. During these years, we have focused on the logistics sector with a product able to monitor and track different assets moving through the ports facilities such as containers, trains, trucks, etc. We are working today with the major Spanish ports and have

clients in several other countries such as France, Italy, Portugal, Belgium, Switzerland, Brasil, Mexico and Colombia.

The MSc project will be developed within our existing research line on access point monitoring in logistic environments, and the student will work closely with the Deep Learning research and technical team of AllRead.

PAID PROJECT AND FUTURE PERSPECTIVES
In parallel to the Master project we offer the possibility to do practicals at AllRead, which will offer the student the possibility to dedicate more time on the project, and extend it to specific application domains.

Both the Master project and the practicals at the company can be paid upon the student experience. High performing students will be offered the possibility to continue professionally as computer vision engineers in AllRead.

Please, do not hesitate to contact us for more clarifications.
The sooner you contact us, the better!

Marçal: marcal@allread.ai
Dimos: dimos@cvc.uab.es