

Cloud-based simulation of page curling in the copying of documents

Fortissimo Experiment Facts:

- Industry Sector: **Digital, Data Analytics**
- Country: **Greece**
- Software Used: **CURLO**

ORGANISATIONS INVOLVED

Bookscanner is a Greek company which specializes in digitising printed material.

Vertoyo, also from Greece, is an SME which specializes in technology development and software service provisioning in the digitization field.

HPC expertise was provided by The Laboratory of Robotics and Automation, Democritus University of Thrace.

Arctur, the Slovenian HPC centre, was the HPC Provider.



THE CHALLENGE

Digitisation of books is an important process, both for commercialisation and for preservation of older texts. The Bookscanner© product can automatically and physically turn pages and automate the scanning process. However, this process results in a 'page curling' effect where the pages are attached to the spine. The scanned pages need to be digitally flattened, a tedious and expensive process that this experiment aimed at improving and commercializing.



THE SOLUTION

A Deep Neural Network (DNN) was trained with simulated page curling of 1 million images. The algorithm takes two already cropped pages from a book image and outputs an artificially curled book page. Using the artificially curled book pages, a state-of-the-art deep Convolutional Encoder-Decoder (CED) Neural Network was trained in order to apply the de-curling process. After training, the DNN can de-curl newly scanned pages with very good success rate. Evaluation of the page de-curling problem showed accuracy of curling correction in over 90% in most cases. With the computational power provided by HPC, the training procedure of the DNN was at least 30 times faster than using a typical workstation.



BUSINESS IMPACT

The only current method for page curling correction is based on a projected laser grid that requires each page to be scanned twice. The solution (called CURLO) removes the need for additional laser grid projection equipment and provides a 50% improvement on the standard curling correction procedure.

As a result of this experiment, the CURLO solution can be offered as a post-processing service to accompany the Bookscanner© product. The collaboration with Arctur has allowed an improvement in the quality of batch-mode scanning. This will be offered as a Software as a Service (SaaS) framework for scanned page de-curling.

Fortissimo Experiment Partners:

- **Bookscanner S.A** (End User)
- **VERTOYO O.E** (ISV)
- **Laboratory of Robotics and Automation, Democritus University of Thrace** (HPC Expert)
- **Arctur** (HPC Provider & Host Centre)

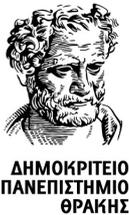
More Information:

www.fortissimo-project.eu
info@fortissimo-project.eu

In addition to the digital content market, the partners in this experiment are ready to address the digitisation needs of a paperless economy, e.g. insurance and paperless banking. Especially regarding the banking sector, recent studies estimate that productivity improves up to 39% when electronic forms replace paper and workflows are used to streamline processes.

BENEFITS

- CURLO has been added to the Bookscanner software, allowing automatic de-curling of pages, saving time compared to the previous method
- This is a scalable solution thanks to the use of HPC and can provide many scanning application domains with automatic capabilities for the amelioration of digital artefacts introduced during the scanning process.
- This is the first such service to be offered via the Fortissimo Marketplace, paving the road to further business development opportunities.
- The monetary benefits for the next 3-5 years can lead to an increase of 10-15% in revenues for participating SMEs. In this case, the payback period for the CURLO investments is 3-4 years. This excess revenue is estimated at approximately €300,000 - €350,000 by the end of year 3 after the end of the project.



THE FORTISSIMO PROJECT

Fortissimo is a collaborative project that enables European SMEs to be more competitive globally through the use of simulation services running on a High Performance Computing cloud infrastructure. The project is coordinated by the University of Edinburgh and involves more than 100 partners including Manufacturing Companies, Application Developers, Domain Experts, IT Solution Providers and HPC Cloud Service Providers from 14 countries. These partners are engaged in over 90 experiments (case studies) where business relevant simulations of industrial processes are implemented and evaluated. The project is funded by the European Commission within the 7th Framework Programme and Horizon 2020 and is part of the I4MS Initiative.

I4MS Fortissimo is part of I4MS ICT Innovation for Manufacturing SMEs: www.i4ms.eu



This project has received funding from the European Union Seventh Framework Programme under grant agreement No 609029 and from the European Union's Horizon 2020 research and innovation programme under grant agreement No 680481.