French name transcription from handwriting using deep learning

1 Description

Handwritten text identification from images is one the most common application domains of deep neural networks [1, 2]. Perhaps the best known dataset used for evaluating deep learning algorithms is the The MNIST Database of Handwritten Digits `yann.lecun.com/exdb/mnist/`. However, more complex datasets of handwritten text have been introduced.

2 Objectives

The goal of the project is to solve the task of name transcription from handwritting images implementing a deep learning approach and using a database with a large number of images of handwritten names[1]. The deep learning approach should be evaluated according to its effectiveness to transcript the written name.

The student should design any preprocessing of the images in the dataset as well as designing the validation approach to evaluate the accuracy of the proposed deep learning approach. As in other projects, a report should describe the characteristics of the design, implementation, and results. A Jupyter notebook should include calls to the implemented function that illustrate the way it works.

3 Suggestions

- Read previous approaches to handwriting transcription using DNNs.
- Any of the available deep network architectures, or features, could be used for the task.
- Implementations can use any Python library.
- If you decide to train a new architecture using the data, provide the trained network for the project revision process. This network can be loaded and invoked from the python notebook.
- Visualization of the features learned by the network is encouraged as an additional step after classification.

References

[1] Dan Claudiu Cireşan, Ueli Meier, Luca Maria Gambardella, and Jürgen Schmidhuber. Deep, big, simple neural nets for handwritten digit recognition. *Neural computation*, 22(12):3207–3220, 2010.

[2] Geoffrey E Hinton, Simon Osindero, and Yee-Whye Teh. A fast learning algorithm for deep belief nets. *Neural computation*, 18(7):1527–1554, 2006.

[1] The Database "Transcriptions of names from handwriting" can be downloaded from `https://www.crowdflower.com/data-for-everyone/`.