A Unified Neural Based Model for Structured Output Problems

Soufiane Belharbi*† , Clément Chatelain*‡[1], Romain Hérault*§ , and Sébastien Adam*¶[1]

[1]das2014 – das2014 – France

Résumé

Structured output problems are characterized by structural dependencies between the outputs (e.g. the classes distribution in image labeling problem, the words positions in sequence tagging in natural language processing). Traditionally, graphical models such as HMM and CRF are used to capture the interdependencies of the outputs. In this article, we propose a unified framework to deal with this problem where we combine learning the hidden interdependencies of the inputs and the outputs in the same optimization. In our framework, we extend the input pre-training layer technique for deep neural networks to pre-train the output layers aiming at learning the outputs structure. We propose a neural based model, called Input/Output Deep Architecture (IODA) to solve the optimization.
Facial landmark detection is a real-world application where the output key points of the face shape have an obvious geometric structure dependencies. We perform an evaluation of IODA on this task over two challenging datasets: LFPW and HELEN. We demonstrate that IODA outperforms a deep network with the traditional pre-training technique.

Mots-Clés: structured output data, deep learning, auto, encoders, stacked auto, encoders, facial landmark detection

*Intervenant
†Auteur correspondant: soufiane.belharbi@litislab.fr
‡Auteur correspondant: clement.chatelain@insa-rouen.fr
§Auteur correspondant: romain.herault@insa-rouen.fr
¶Auteur correspondant: Sebastien.Adam@litislab.eu