



HAL
open science

Recognition and segmentation of children handwriting

Simon Corbillé, Elisa Fromont, Eric Anquetil

► **To cite this version:**

Simon Corbillé, Elisa Fromont, Eric Anquetil. Recognition and segmentation of children handwriting. SIFED 2022 - Symposium International Francophone sur l'Écrit et le Document, Oct 2022, Rennes, France. hal-03879307

HAL Id: hal-03879307

<https://hal.science/hal-03879307v1>

Submitted on 30 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

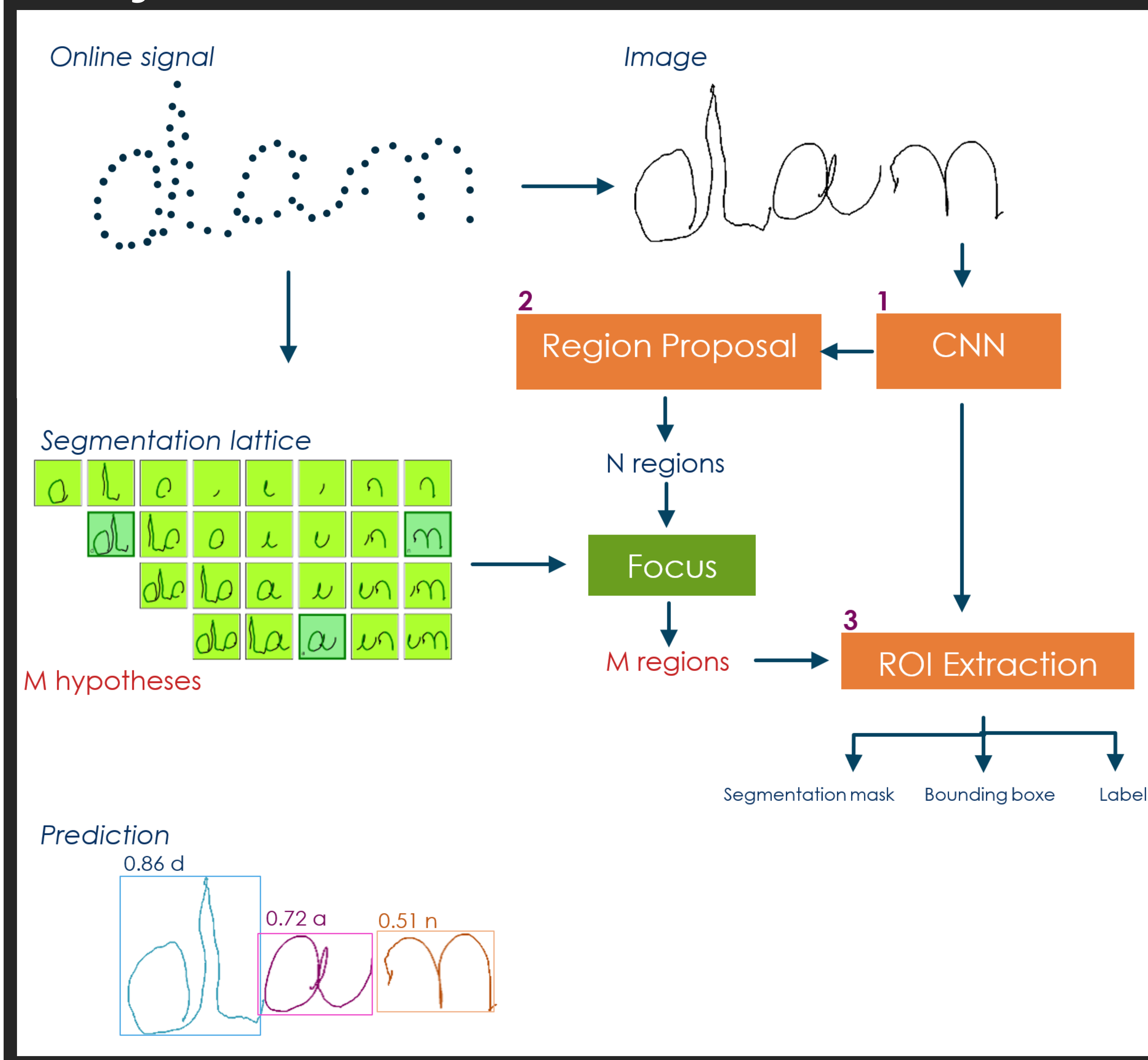
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Recognition and segmentation of children handwriting

Simon Corbillé, Élisabeth Fromont, Éric Anquetil,
Univ Rennes, IRISA lab 35000 Rennes France

Contact : simon.corbille@irisa.fr

Use segmentation lattice to focus Mask R CNN:



DATASET

The dataset is composed of children words annotated on letter level. The handwriting is encoded with online format. There are converted to image resize at 128px height and pad at 1280px width.

Data augmentation is performed on train part on 375 words to obtain 12 000 words with:

- Offline transformations :
 - Rotation, inclination, stretching.
- Online transformation :
 - Curvature, stroke stretching.

Split

Train	Validation	Test
12000	128	1032

PRELIMINARY RESULTS

Evaluation metrics :

- Character Error Rate (CER) : evaluate the recognition aspect
- Intersection Over Union (IoU) : evaluate the segmentation aspect

Methods	CER (%)	IoU (%)
Baseline	14,4	78,1
With lattice	13,0	77,6

The use of segmentation lattice improve the **CER by 1,4%** and decreases slightly the IoU by 0,5%. The preliminary results are encouraging, but the recognition performance is not acceptable for a real application use. More experiments on how to use the lattice are in progress.

CONCLUSION AND PERSPECTIVES

The contribution of this work is to use a segmentation lattice with a Mask R CNN. The first results show the potential of this approach, although the recognition still not efficient enough for a real application.

Another idea is to use a hybridization of a model that is great in recognition and approximate in segmentation (Seq2Seq) and of a model that is approximate in recognition and efficient in segmentation (Mask R CNN).

INTRODUCTION

Handwriting recognition is a challenging, already well-studied topic. However, most research focus on adult handwriting recognition. Few are interested on children handwriting and to provide a **character segmentation** associate with the **recognition**. Moreover, **children handwriting** is a more difficult context because children are in an early stage of learning. Their handwriting is **deformed** and contain **spelling mistakes**.

The goal of this work is to design an efficient model to do recognition and segmentation of children handwriting for an education purpose.

METHOD

The first step is to select a model which can do recognition and provide an accurate segmentation. We chose an **object based detection** architecture : **Mask R CNN**. The main steps of this model are :

1. **Extract features** from input image with a Convolutional Neural Network (CNN).
2. **Compute coarse regions** to analyze with a Region Proposal Network (RPN). Because the network do not know the size of object to detect, it uses a lot of regions with different scales, ratios and positions. For each region, RPN predict an objectness score and bounding box position.
3. **Extract features for each region** and predict the label, bounding boxes and segmentation mask. Here, the bounding boxes prediction makes it possible to refine the delimitation proposed by the RPN.
4. **Non-maximum Suppression (NMS)** is applied to select the best predictions.

The next step is to improve the model with expert knowledge. The handwriting is encoded with an **online signal converted to an image** as input of Mask R CNN. Using knowledge relative to handwriting, the online signal can be segmented into different hypotheses of letter regroup in a **segmentation lattice**. The goal is to use the lattice to focus the model on precise regions of the image in RPN.