platform{DH}, December 4

Machine vision algorithms on cadaster maps

Sofia Ares Oliveira

Ecole Polytechnique Fédérale de Lausanne
Digital Humanities Laboratory

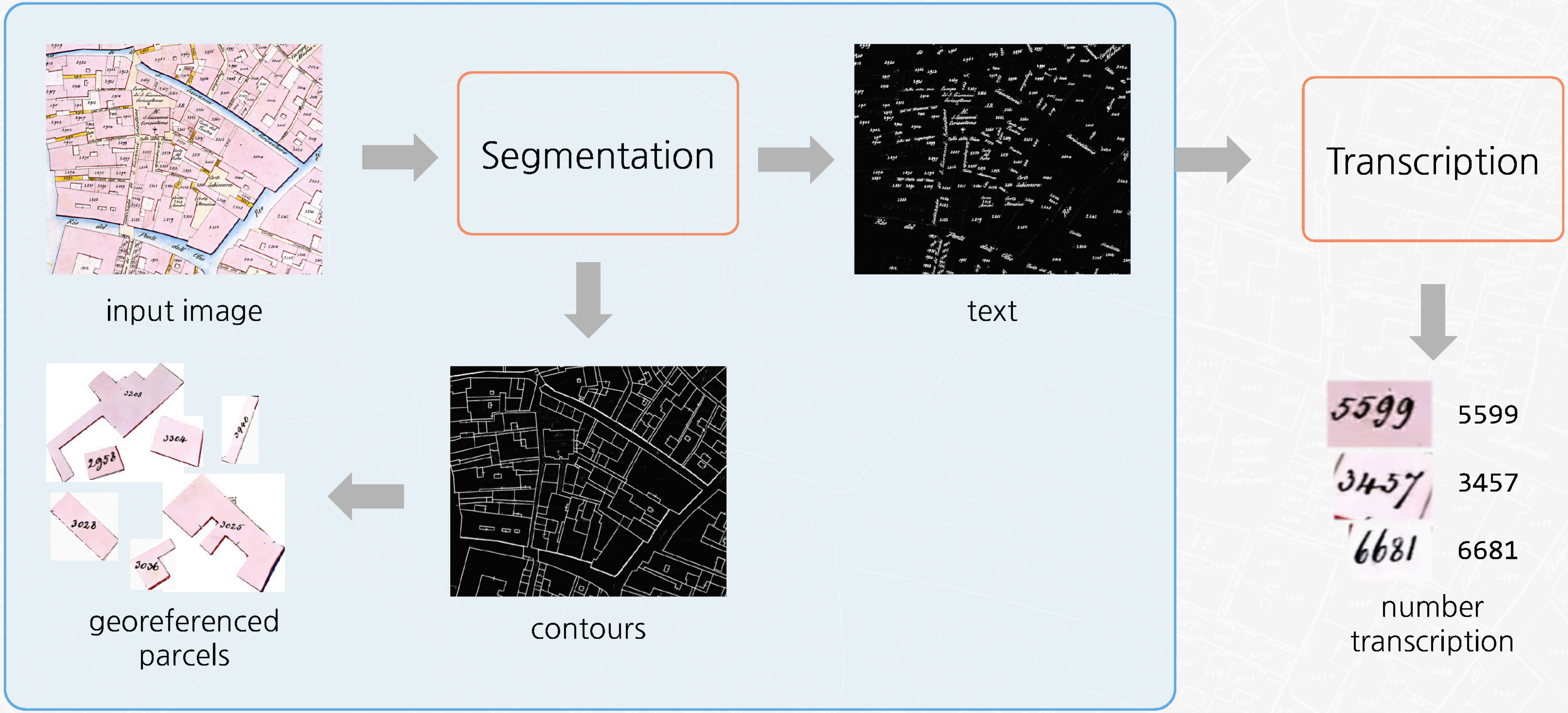
Venice
1808

~~Comune~~ Città di Venezia — Dipartimento dell'Adriatico

Numeri della Mappa	subalterni	POSSESSORI	Denominazione dei Pezzi di terra	QUALITA'	Classe	SUPERFICIE Pertiche Censuarie	Centesimi
			Ponte di Rialto				87
9001		Grimani . . . . . . qm	30	Bottega d'Affitto			
9002		Marcello Marino qm Guido	29 C.S.	Simile			
9003		Venier Prò Gerolamo qm Gio: Batta	28 C.S.	Simile			
9004		Suddetto	27 C.S.	Simile			
9005		Marconi Giuseppe qm Ant.o	26 C.S.	Bottega di proprio uso			
9006		Raspi Fran.co qm Gio: Maria	25 C.S.	Bottega d'Affitto			
9007		Pesenti Andrea qm	4381 C.S.	Simile			
9008		Morosini Ant.o di Vincenzo Dolfin Luigi qm . . . . Dolfin Marcant.o qm . . . . Ferracina Niccolò qm . . . . Petrogalli Gius.e qm . . . . Bergonti Gius.e qm . . . . Dolfin Marianna qm . . . . Possessori indivisi	4580 C.S.	Simile			
9009		Marcuzzi Costantino qm Zanno	4579 C.S.	Simile			
9010		Dolfin Leonardo qm Prò	4578 C.S.	Simile			
9011		Venier Gerolamo qm Gio: Batta	4577 C.S.	Simile			
9012		Eredi del fu Paolo Emilio Canal qm Gerolamo. Balbi Canal Mattia Ant.a Vedova del fu Cristoforo Possessori indivisi	4577 C.S.	Simile			
9013		Marcuzzi Costantino qm Zanno	4587 C.S.	Simile			
9014		Correr Marcello Maria qm Prò	4586 C.S.	Simile			
9015		Labia Fran.co qm Paolo Ant.o	4585 C.S.	Simile			
9016		Venier Prò Gerolamo qm Gio: B.a	4584 C.S.	Simile			

Method

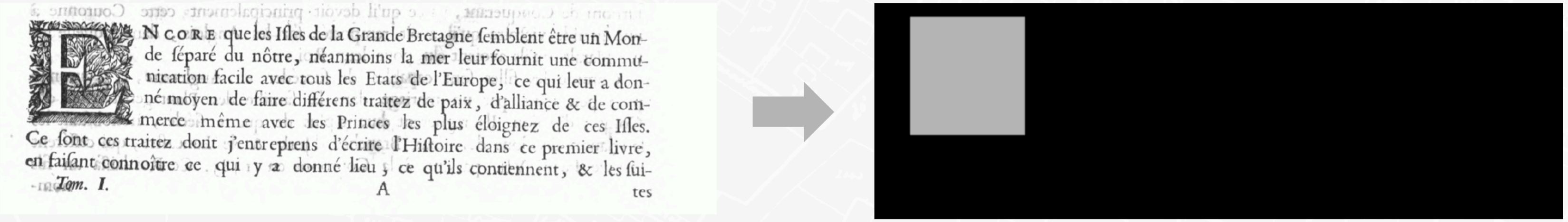
Overview of the system



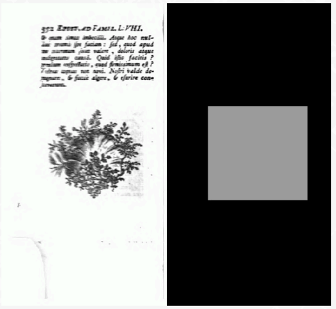
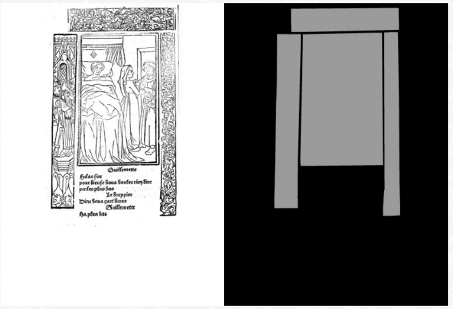
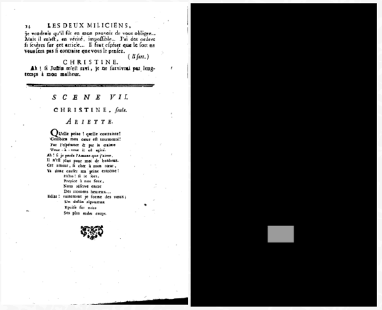
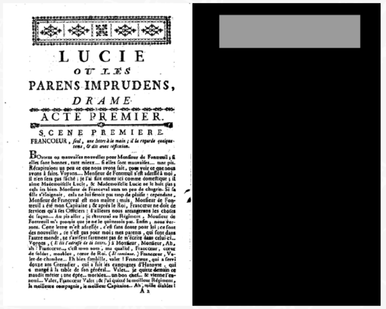
Segmentation

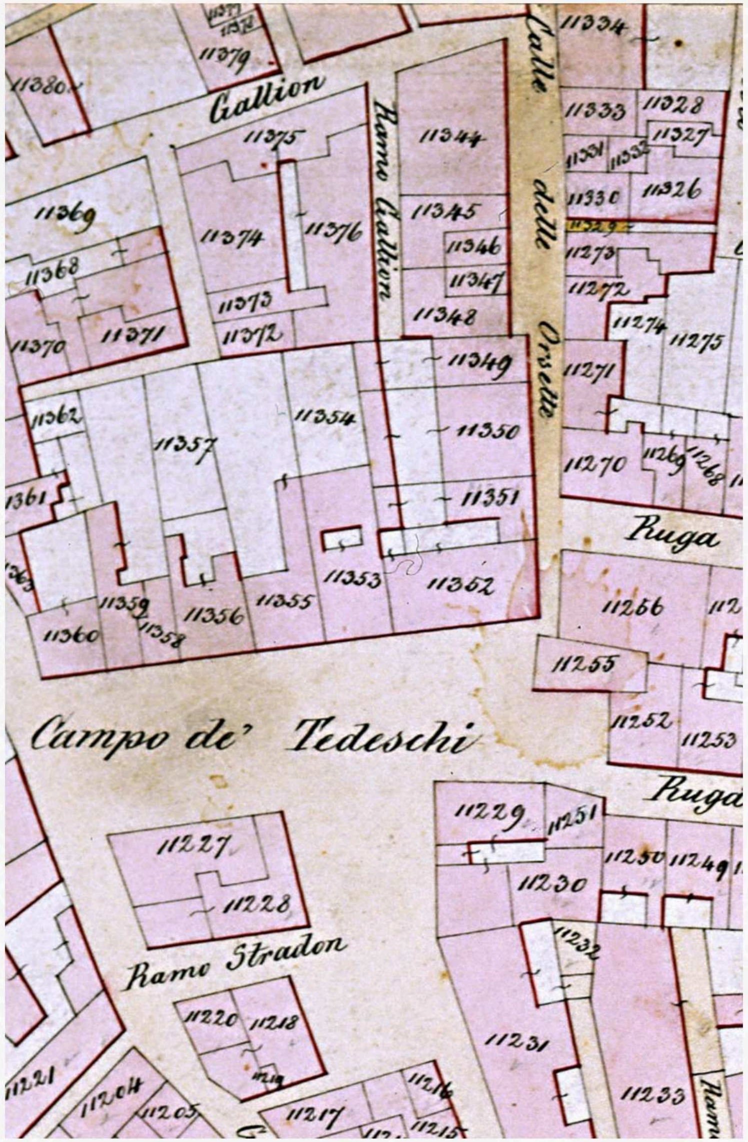
Deep learning pixel-wise segmentation

Assign to each pixel a label

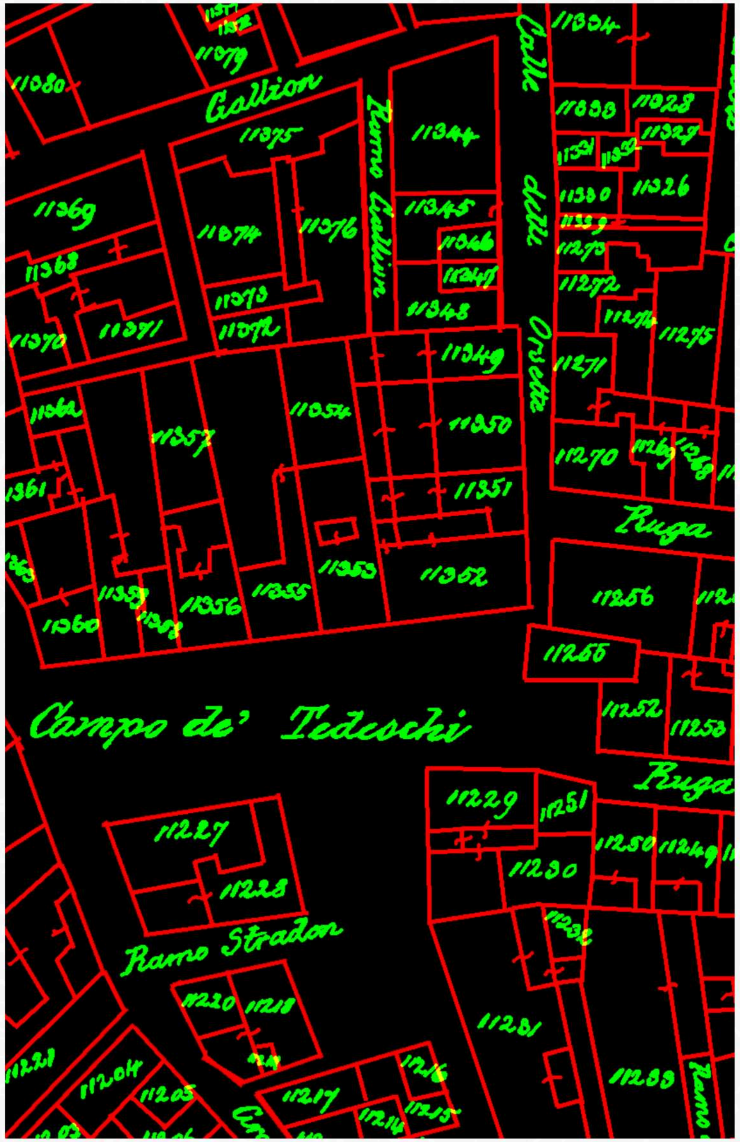


Train a neuronal network by showing several examples



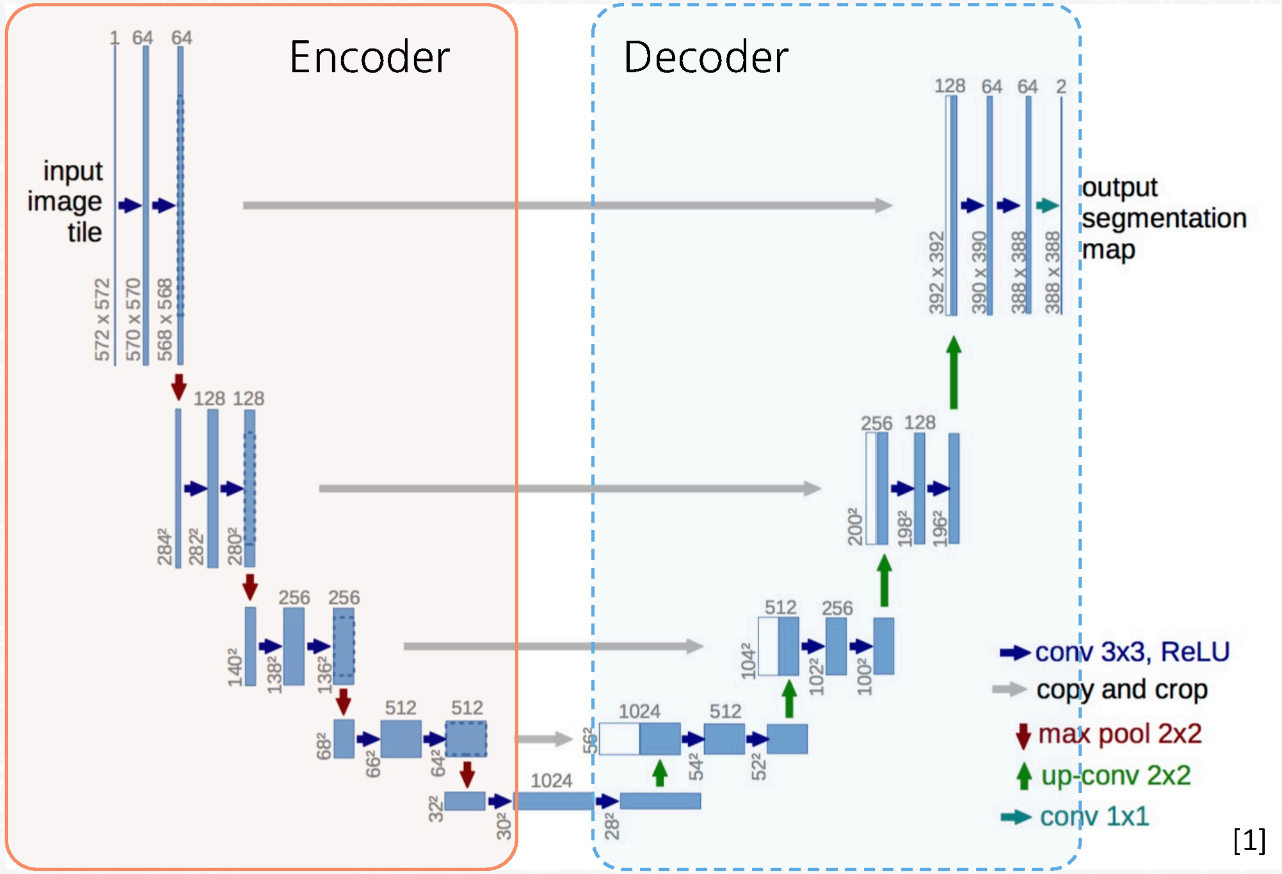
Data annotation
Pixel labelling



Architecture network

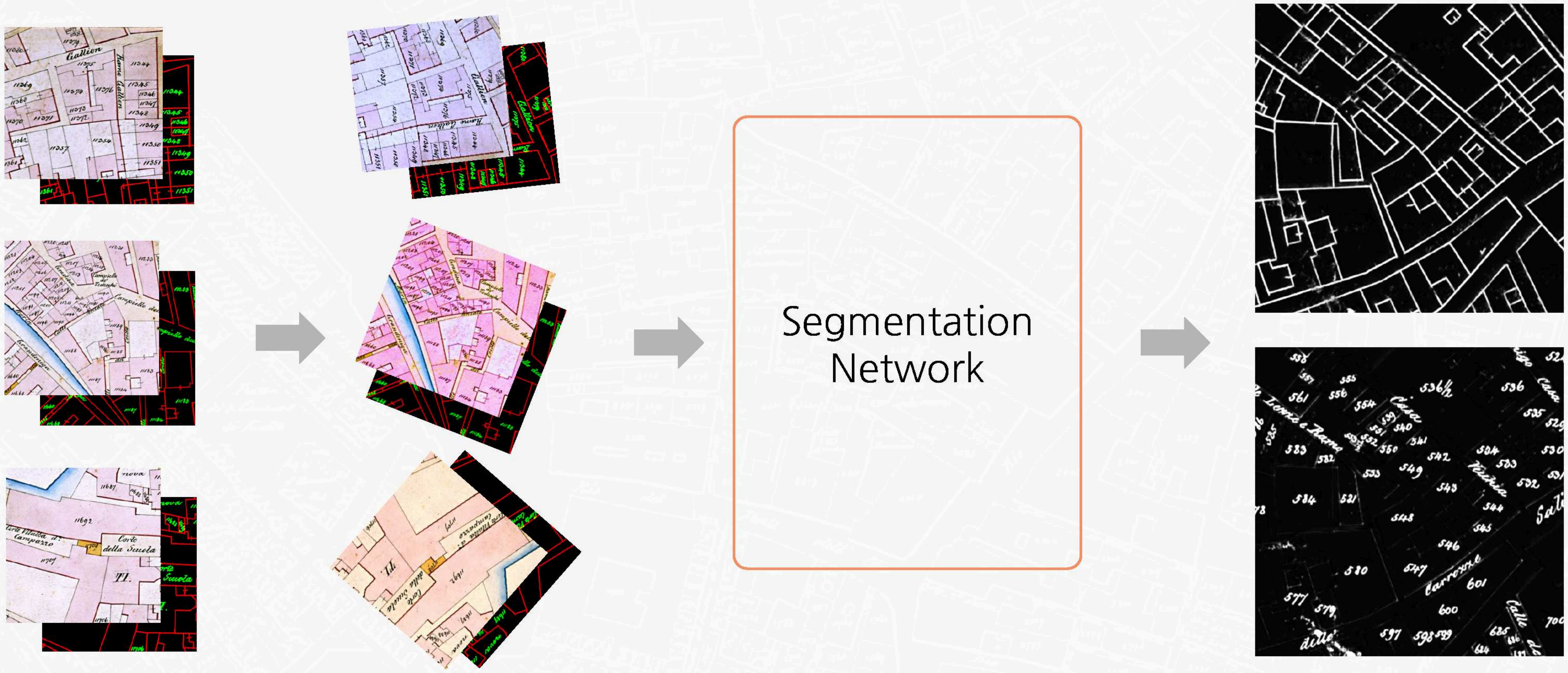
- Encoder is a pre-trained convolutional neural network
- Each encoding layer has a corresponding decoding layer

➔ U-shaped architecture



[1] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation", 2015
[2] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation", 2015

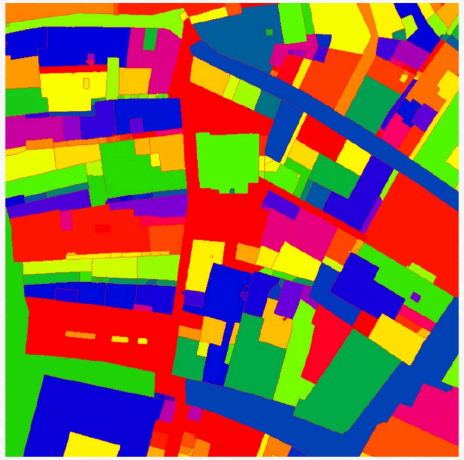
Training



Input {image, label}   data augmentation   neural network   2-class prediction

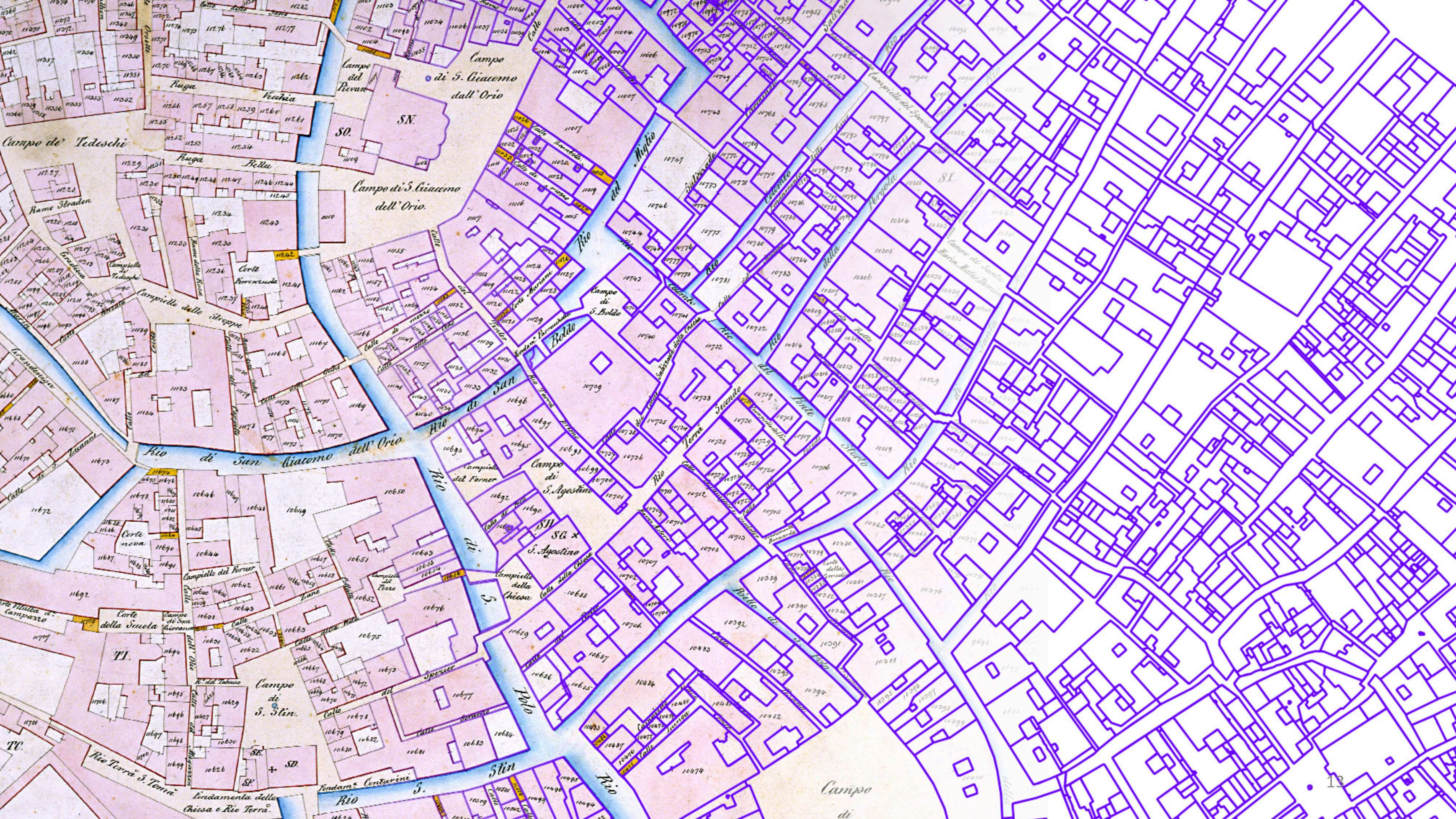
Parcel extraction and georeferencing

watershed



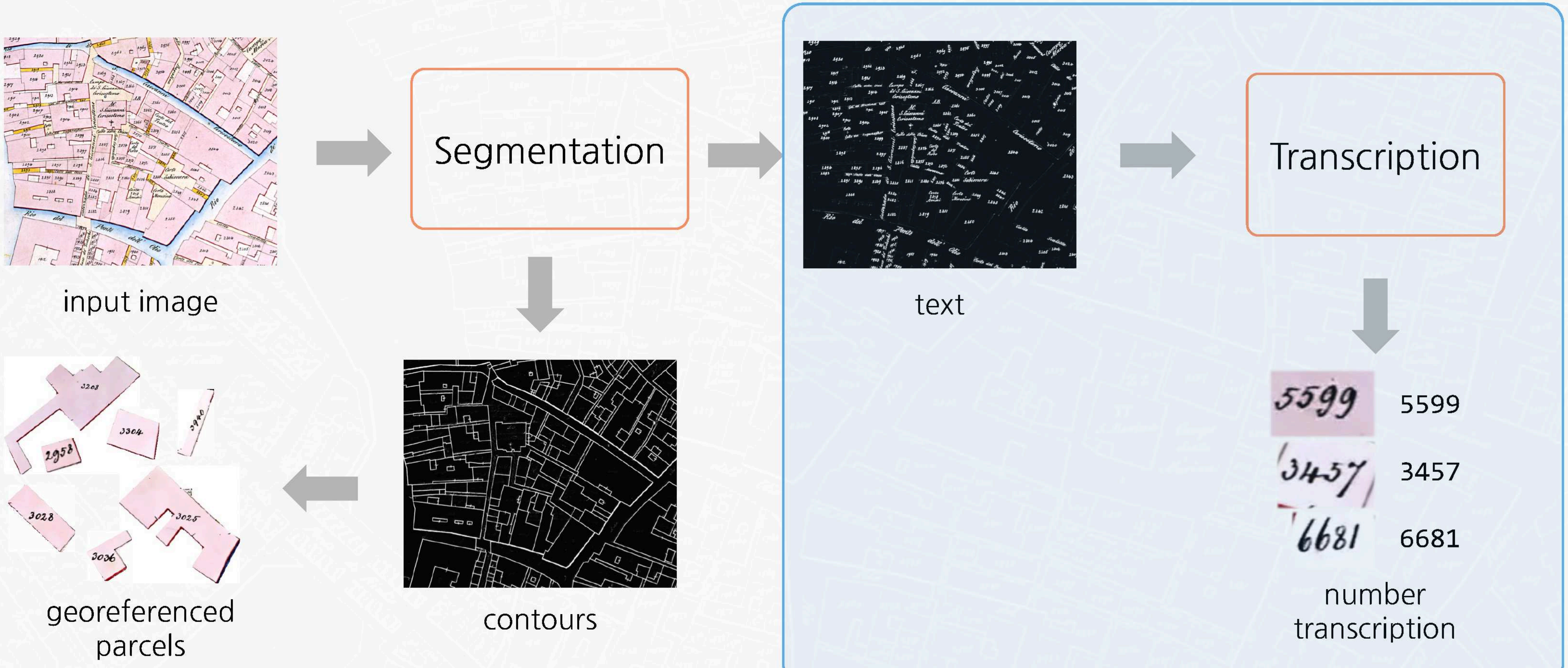
contour extraction

If the image is georeferenced, the geographical coordinates are directly inferred and the parcels can be exported into a GIS system


Campo de' Tedeschi
Campo di S. Giacomo dall' Orio
Campo del Piovan
Campo di S. Giacomo dell' Orio
Ruga Vecchia
Ruga Bella
Rio di San Giacomo dell' Orio
Rio di S. Boldo
Campo di S. Boldo
Campo di S. Agostino
S. Agostino
Campiello del Forner
Campiello della Chiesa
Corte nova
Corte della Scuola
Campo di S. Stin
Rio di S. Polo
Fondamenta Contarini
Rio Terrà S. Tomà
Rio del Megio
Rio Terrà Secondo
Campo

Campo di S. Giovanni Grisostomo
AC.
S. Giovanni Grisostomo
S. Giovanni Grisostomo
Corte del Teatro
Calle del Teatro
Corte Sabionera
Corte Morosina
Rio del Ponte dell' Olio
Rio di Santa Marina
Campo Santa Marina
YY.
S. Marina
ZZ
Miracoli
Rio
Campo SS. Giovanni e Paolo
Corte del Forno
Calle della Malvasia
Campo San Lio
SS.
RR.
NN.
S. Bartolomeo
Salizzada San Bartolomeo
Salizzada del Fondaco de Tedeschi
Rio del Piombo
Calle del Ponte di S. Antonio
Calle della Bissa
Corte Borella
Borgoloco

Pipeline


input image
Segmentation
text
Transcription
contours
georeferenced parcels
5599
3457
6681
number transcription

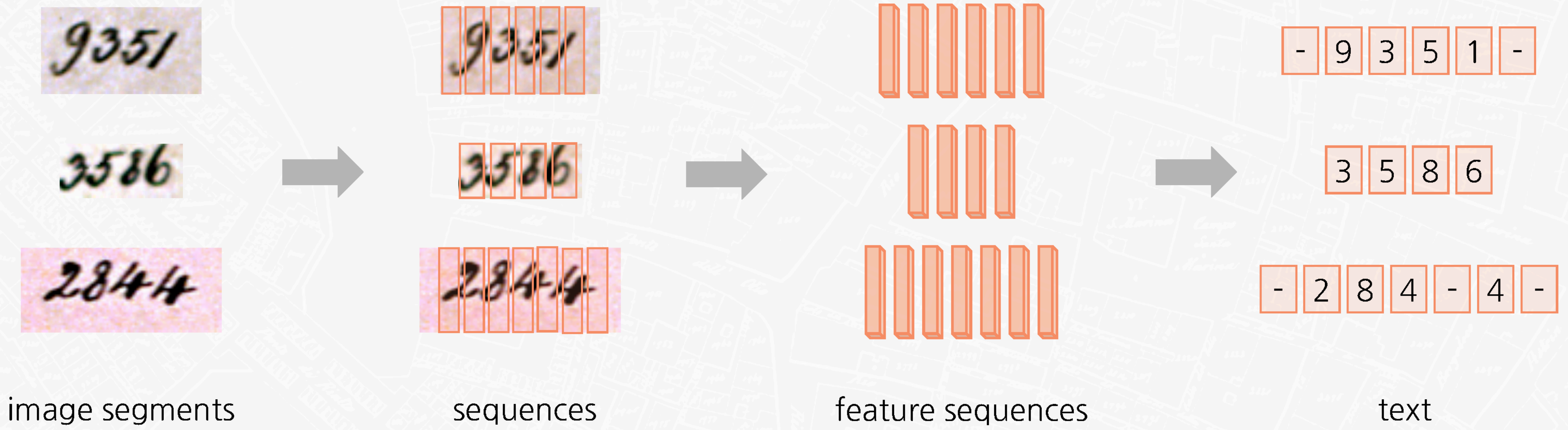
Text extraction

For each parcel, find its corresponding text region

parcel localization | text probabilities | orientation finding | rotation and cropping

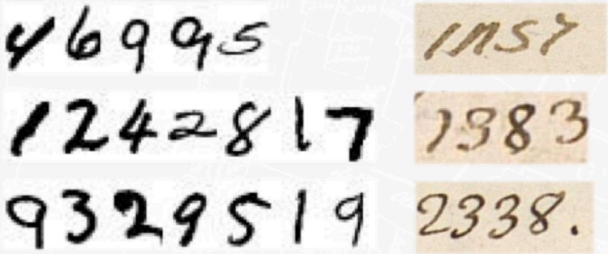
Transcription

Convert image segments into features sequences that will be mapped into text



Training data

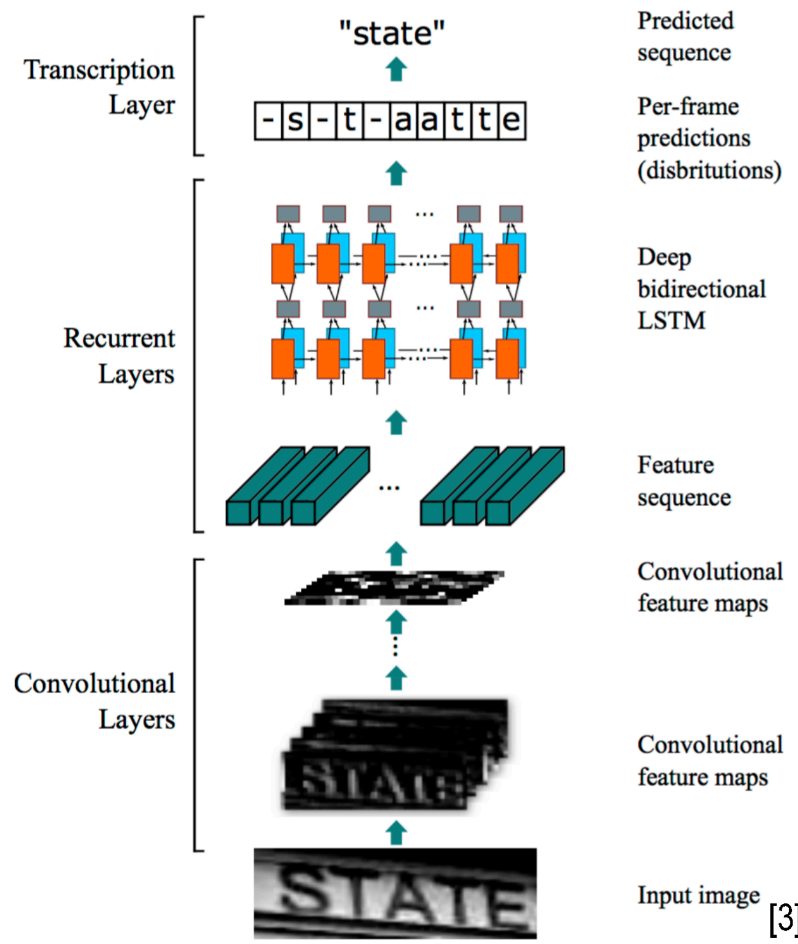
- Synthetic data generated from MNIST dataset (100K)
- Handwritten numbers extracted from venetian archives (~ 30K)



+ data augmentation

Architecture network : CRNN

1. Convolutional neural network (CNN)
2. Recurrent neural network (RNN) with bidirectional Long Short-Term Memory (LSTM)
3. Mapping of separated time step labels to sequence label with connectionist temporal classification (CTC)



[3]

[3] B. Shi et al. "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," 2017
[4] A. Graves, et al. "Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks,", 2016

Results

Parcel extraction results

IoU threshold	0.7	0.8	0.9
Parcel recall	0.90 (1062)	**0.79** (941)	0.51 (605)
Parcel precision	0.50	0.44	0.28
Extracted parcels	2121		
Ground truth	1185		

$$recall = \frac{true\ positives}{total\ ground\ truth} \in [0, 1]$$

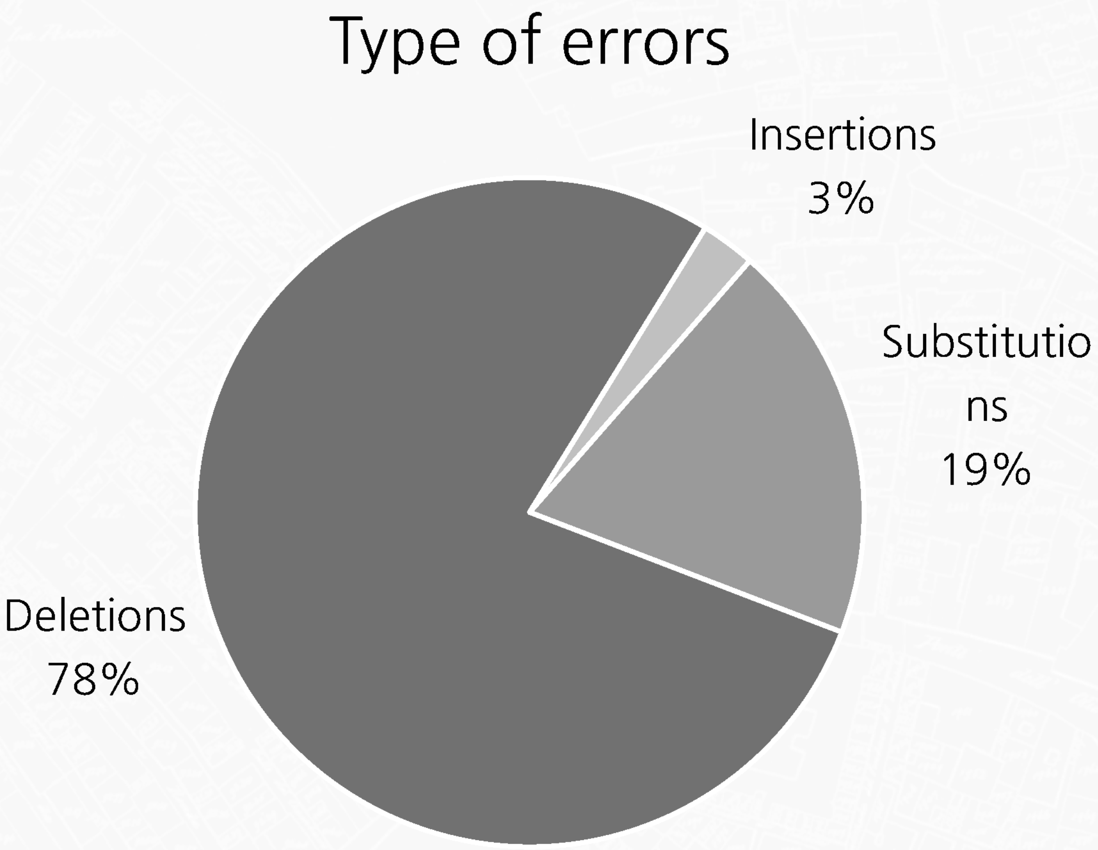
$$precision = \frac{true\ positives}{total\ retrieved} \in [0, 1]$$

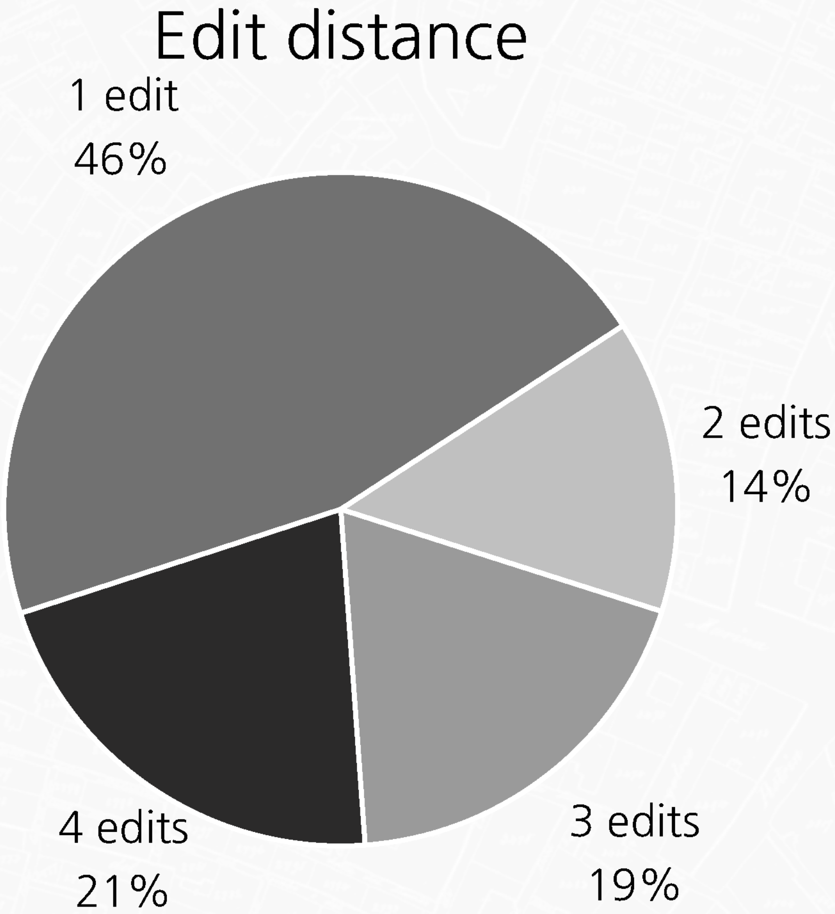
Label extraction and transcription results

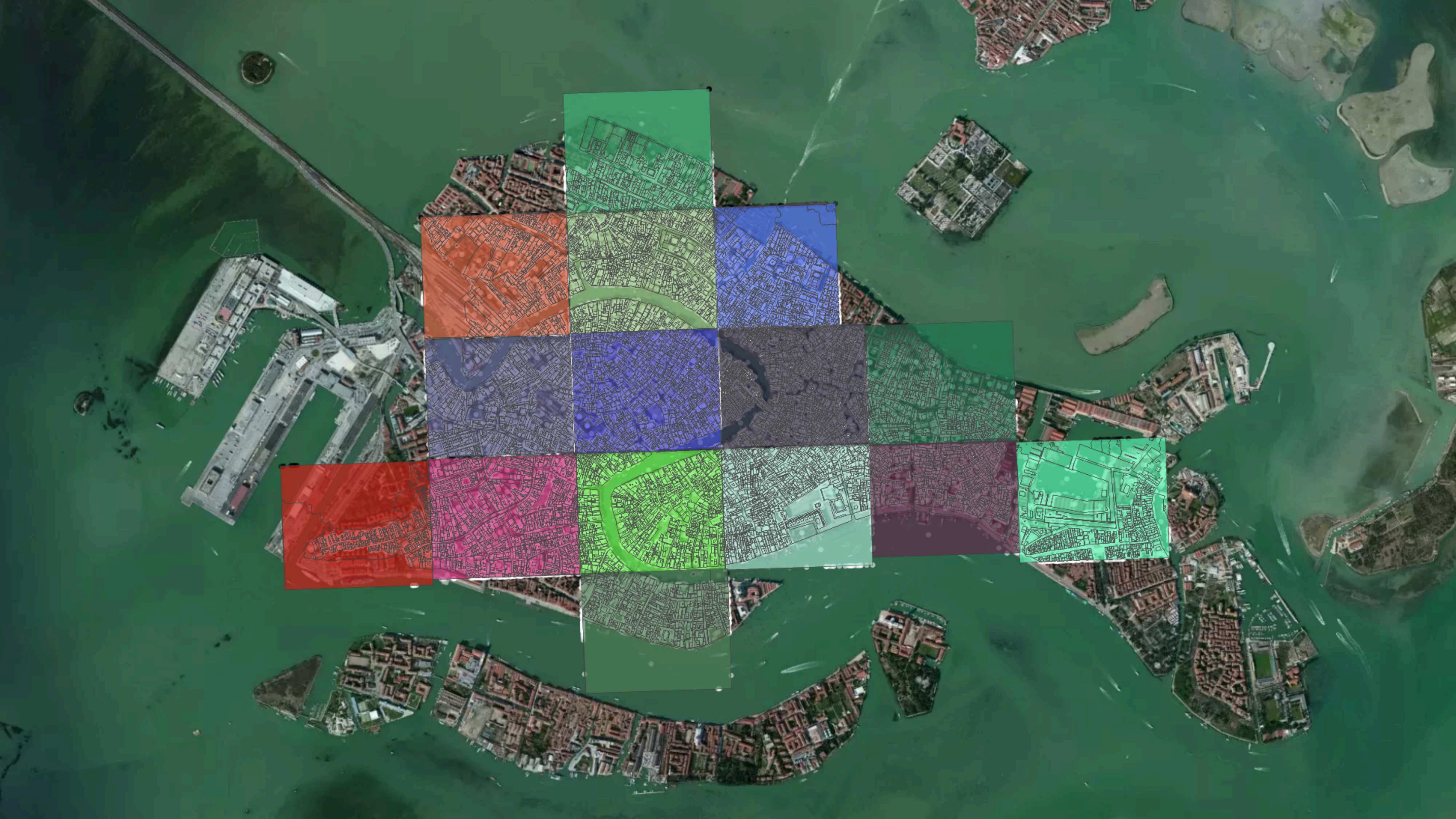
Label locating	Inter : 0.8
Recall	0.86 (633)
Precision	0.37
Extracted labels	1693
Ground truth (labels)	736

Label transcription	Inter : 0.8
Recall (correctly transcribed)	**0.83** (608)
Precision	0.36
Character Error Rate (CER)	**0.14**
Ground truth (labels)	736

Label transcription: what are the errors ?







Conclusion

The system automatically extracts the parcels and their labels with high confidence level and opens new perspectives for spatial analysis in social, economic and urban structures.

Sofia Ares Oliveira
sofia.oliveiraares@epfl.ch

github.com/dhlab-epfl/cadasters
github.com/solivr

Digital Humanities Laboratory
dhlab.epfl.ch

Isabella di Lenardo
isabella.dilenardo@epfl.ch

Frederic Kaplan
frederic.kaplan@epfl.ch

Benoit Seguin
benoit.seguin@epfl.ch

References

1. RONNEBERGER, Olaf, FISCHER, Philipp, et BROX, Thomas. U-net: Convolutional networks for biomedical image segmentation. In : *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2015. p. 234-241.
2. BADRINARAYANAN, Vijay, KENDALL, Alex, et CIPOLLA, Roberto. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv:1511.00561*, 2015.
3. SHI, Baoguang, BAI, Xiang, et YAO, Cong. An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE transactions on pattern analysis and machine intelligence*, 2016.
4. GRAVES, Alex, FERNÁNDEZ, Santiago, GOMEZ, Faustino, *et al.* Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks. In : *Proceedings of the 23rd international conference on Machine learning*. ACM, 2006. p. 369-376.