Towards binarization of Iron Age ostraca from multispectral weakly-annotated imaging

Ohr Dallal[1], Shira Faigenbaum-Golovin[2], Israel Finkelstein[3], Nachum Dershowitz[4]

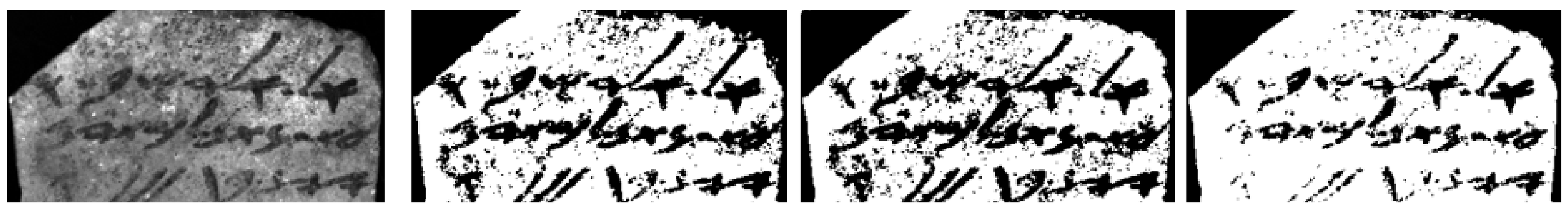
1. School of Electrical Engineering, Tel Aviv University, Israel. 2. Department of Mathematics, Duke University, US. 3. Jacob M. Alkow Department of Archaeology and Ancient Near Eastern Civilizations, Tel Aviv University, Israel. School of Archaeology and Maritime Cultures, University of Haifa, Israel. 4. School of Computer Science, Tel Aviv University, Israel.

Objective

We aim to binarize Iron Age Hebrew ostraca, which are of great importance to the historical study of ancient Judah (dated ca. 600 BCE).

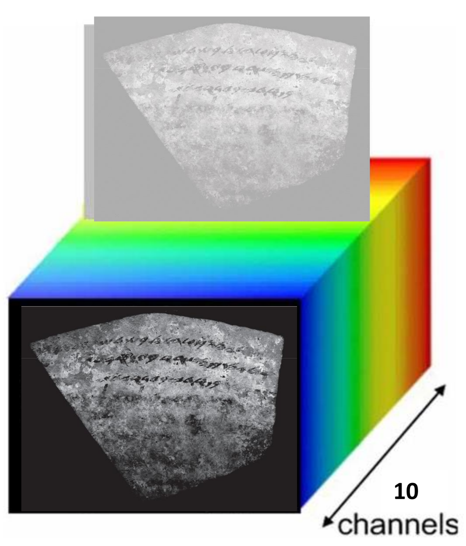
Image binarization is the task of classifying each pixel in an image into either foreground or background, and is one of the essential and preliminary steps towards many document processing tasks.

The working hypothesis is that the multispectral signature of the ink differs from the one of the clay, due to chemical difference.



Methods

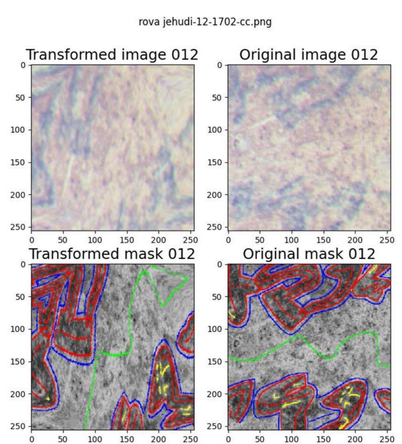
We propose to apply end-to-end deep neural networks (DNN) for exploiting complex shape and spectral cues available in the data, which allow for better discrimination of ink and background and possibly even the reconstruction of ink invisible to the naked-eye. We develop and test DNN binarization models, based on segmentation feedforward neural networks with convolutional and self-attention blocks. The following diagram describes the different methods currently being researched.



A multispectral cube [4]

Dataset split: *Train - Validation - Test*

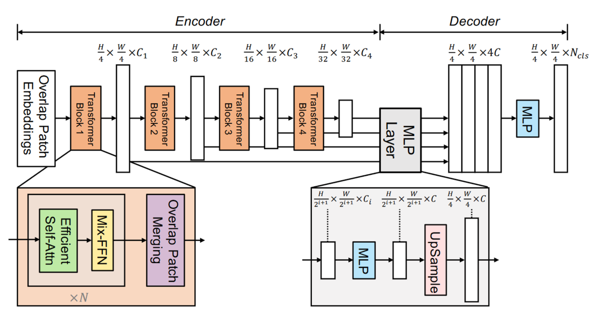
Cross-validated training and evaluation



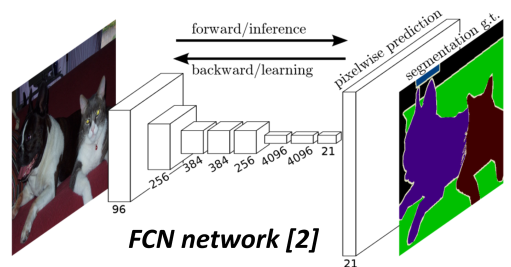
Augmented image and mask

Online Data Pre-processing

- **Split** ostracon's image to random crops
- Data-balanced **sampling** and **normalization**
- Diverse random **augmentation** transforms - generating different representations of the input data (shift, scale, rotate, flip, Elastic transform, etc.)



Transformer-based model architecture [5]



FCN network [2]

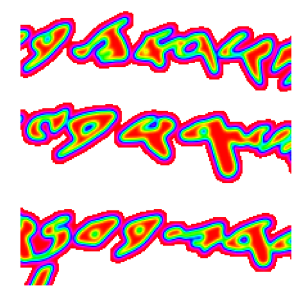
DNN Model

- **Multiscale Features** - combine low-level and high-level spatio-spectral features in neural network's architecture
- Two DNNs designs are being developed:
  - **Convolutional-based – Fully** Convolutional Networks (FCN) [2]
  - **Transformer-based** - with spatial/spectral self-attention for local and global predictive context

Backpropagation

Prediction & Evaluation

- **Tune** hyper-params with Bayesian optimization
- **Stitch** crops for final testing



Loss function

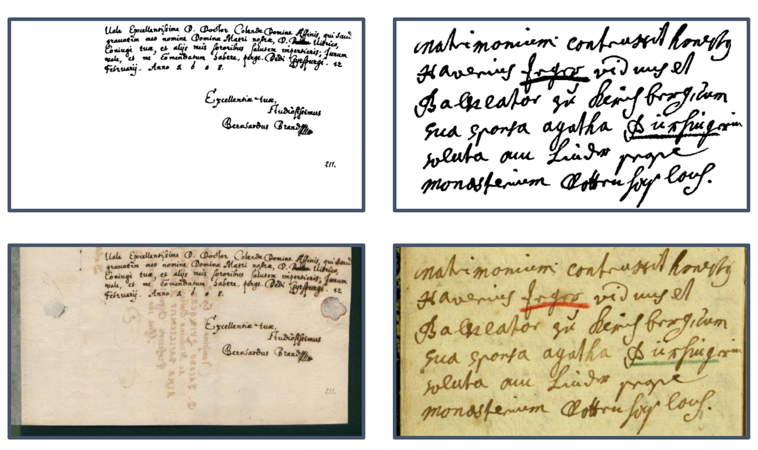
- **Weighted** computation of cross-entropy loss according to each label type
- **Robustness to outliers**
- **Hard/easy examples mining**
- **Regularization** - label smoothing, dropout, etc.

Proposed binarization pipeline. Multiple techniques are used to tackle the problem's complexity.

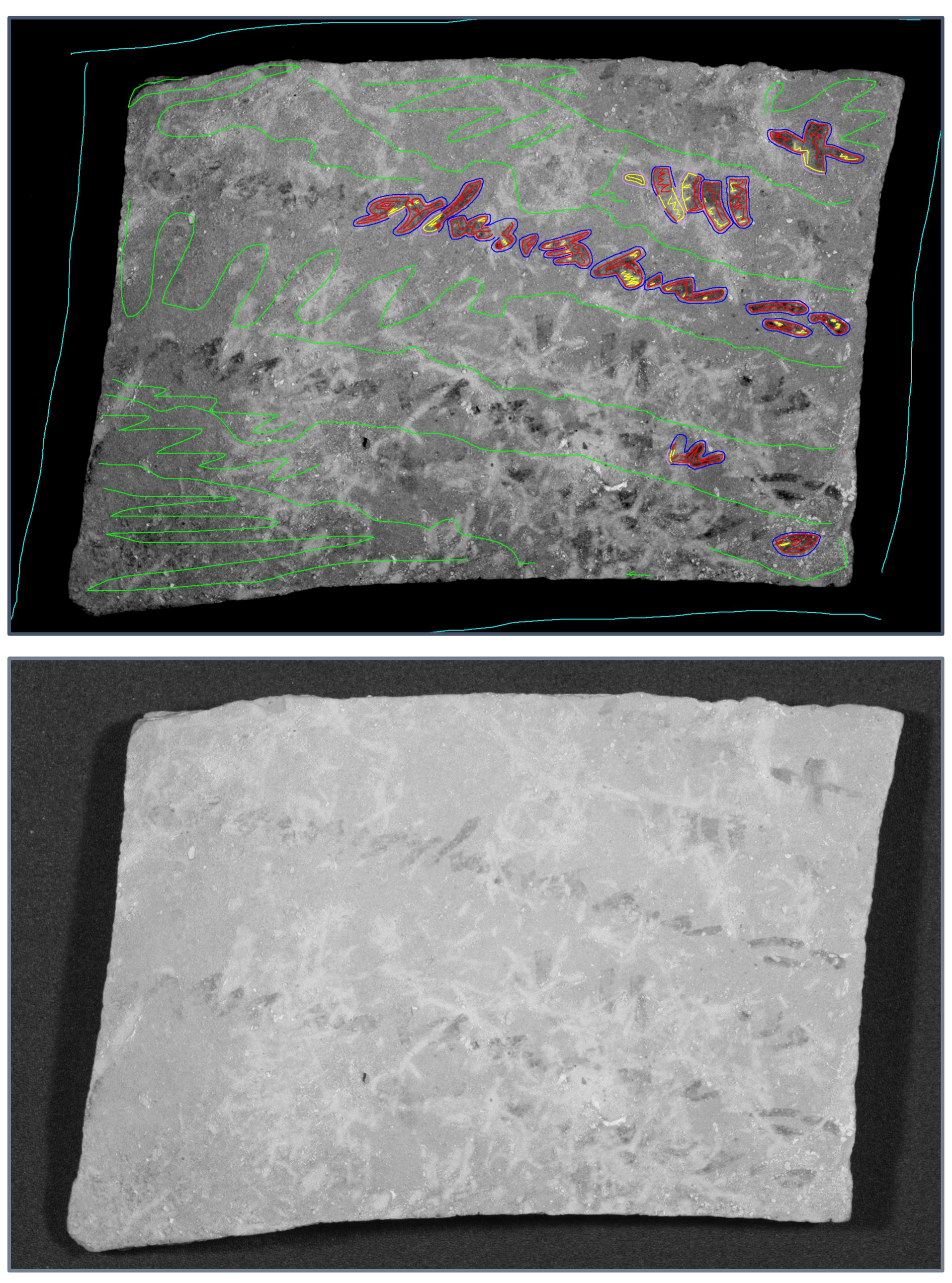
Data

We introduce a new dataset of ancient ostraca images from the First Temple Period manually annotated by human experts. The following comparison demonstrates what makes our dataset unique and challenging:

Properties	Other typical text binarization datasets (e.g. DIBCO and H-DIBCO [1])	Our dataset
Number of samples	Various	Small scale, a total of ~60 ostraca
Number of channels	RGB or grayscale	Multispectral - 10 camera wavelengths
Supervision	Fully-supervised - labels are available for all pixels	Weakly-supervised - sparse and partial scribble labels, possibly equivocal
Complexity	Varying complexity	Severe challenges such as faded text, degradations, contrast variation, etc.
Alphabet	Various handwritten and printed texts	Ancient Hebrew
Utilization of external datasets	Transfer learning is possible	Transfer learning is harder and less useful due to the unique properties of data



H-DIBCO 2018 dataset sample images (bottom) and their ground-truth labels (top) [1]



"Yarmut-12-1785" ostracon from our dataset (bottom) and its ground-truth labels (top)

Our dataset's labels legend:

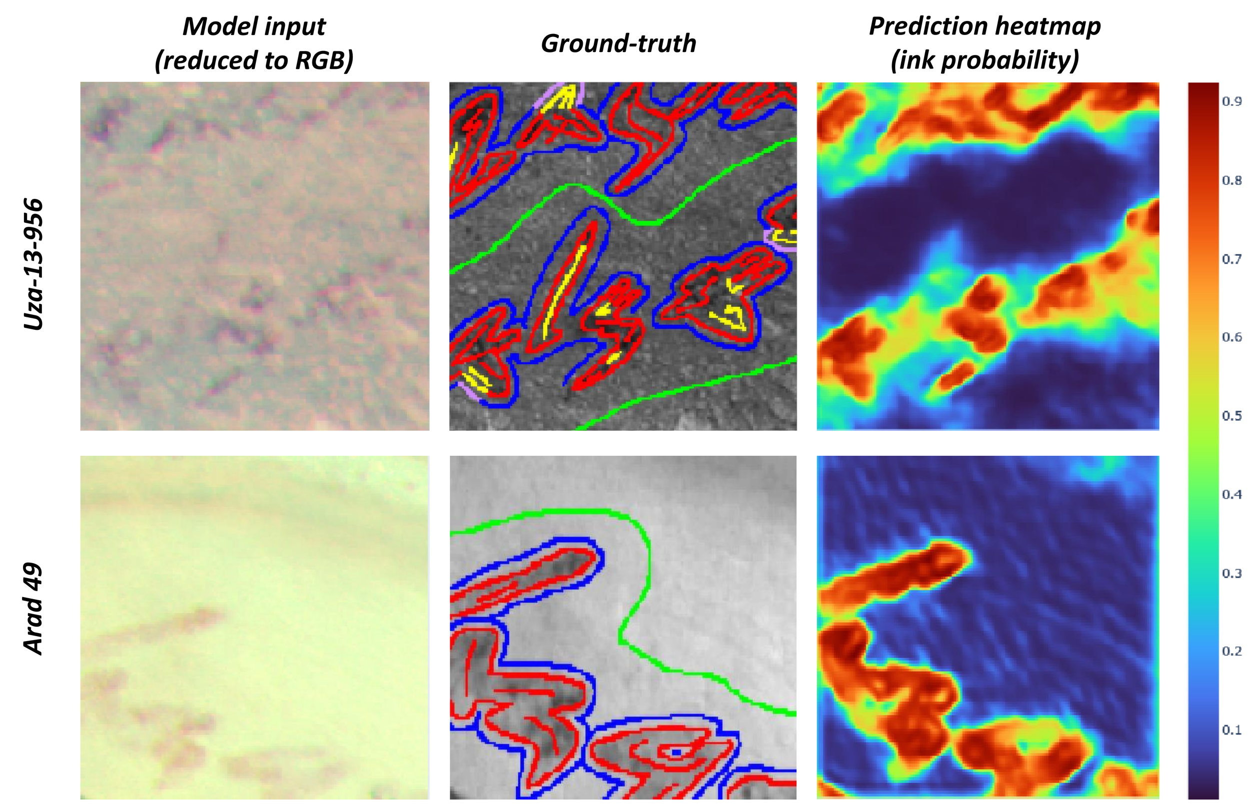
Red = Ink inside letters
Blue = Background outline of visible letters
Green = Random background between letters
Light blue = Background outside the ostracon
Yellow = Ink not / hardly visible but known to exist
Lavender = Background outline of invisible ink
Orange = Pixels of modern text

Camera wavelengths (nm):
525, 590, 635, 660, 695, 735, 775, 830, 890, 940

Ostraca's origins:
From Iron age Judah (ca. 600 BCE) –
Arad, Lachish, Uza, Yarmut, etc.

Current results

Predictions on evaluation set are demonstrated below:



Future work

- **Exploiting unlabeled data** - further researching weakly- and semi-supervised techniques.
- Improving generalization by further developing **special augmentations** (for example, by mixing different training samples).
- **Self-supervision** on auxiliary tasks for dealing with the scarcity of data and learning useful data representations before the binarization task.
- **Labeling** – proceeding the manual ostraca annotation effort.
- Continuing development and optimization of **models and training methods** (e.g., Curriculum Learning, spectral / spatial attention blocks).
- **Performance analysis** and possibly applying the method to additional datasets.

References

[1] Pratikakis, Ioannis et al. "ICFHR 2018 Competition on Handwritten Document Image Binarization (H-DIBCO 2018)." 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR) (2018): 489-493.
[2] Shelhamer, Evan et al. "Fully convolutional networks for semantic segmentation." 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2014): 3431-3440.
[3] Shaus, Arie et al. "Facsimile creation: Review of algorithmic approaches." Alphabets, Texts and Artefacts in the Ancient Near East, Studies Presented to Benjamin Sass (Eds. I. Finkelstein, C. Robin, and T. Römer), Van Dieren, Paris (2016): 474–88.
[4] Sober, Barak et al. "Multispectral imaging as a tool for enhancing the reading of ostraca." Palestine Exploration Quarterly 146 (2014): 185 - 197.
[5] Xie, Enze et al. "SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers." Neural Information Processing Systems (2021).