ARCADA

ARCADA UNIVERSITY OF APPLIED SCIENCES

Handwriting features based detection of fake signatures

29.6.2021

Anton Akusok, Leonardo Espinosa-Leal, Kaj-Mikael Björk, Renjie Hu, Amaury Lendasse

ARCADA

Handwritten signatures

- Handwritten signatures are used daily
- Digital documents are signed with handwritten signatures
- Signature verification is an important practical field of research

Handwritten signature verification

Online signature verification

- Lower error rate
- Needs capture device

Offline signature verification

- Image-based analysis
- Applies to any data format, physical signed papers

Challenge: General signature verification

General signature verification

- Novel research question considered in this work
- Classify genuine/forged signature independently of user, and of true user signature
- Idea of detecting fast automatic hand movement in writing own signature vs slower conscious hand control in forging someone else's signature

Our approach: Local signature features

- MCYT-75 dataset stores signatures as high-resolution images of size 850x360 pixels
- Analysis starts by cutting out small square patches of signature images with sliding window approach
- Image patches show local signature features
- Model based on these features forced to verify signatures based on their writing style, not the shape of full signature

Our approach: Local signature features

- Patch sizes are 256 pixels, 128 pixels and 64 pixels
- Sliding window overlap is 10%, 50% or 90%
- One signature with 90% overlap has ~100 large patches, ~1000 medium patches, and ~5000 small patches

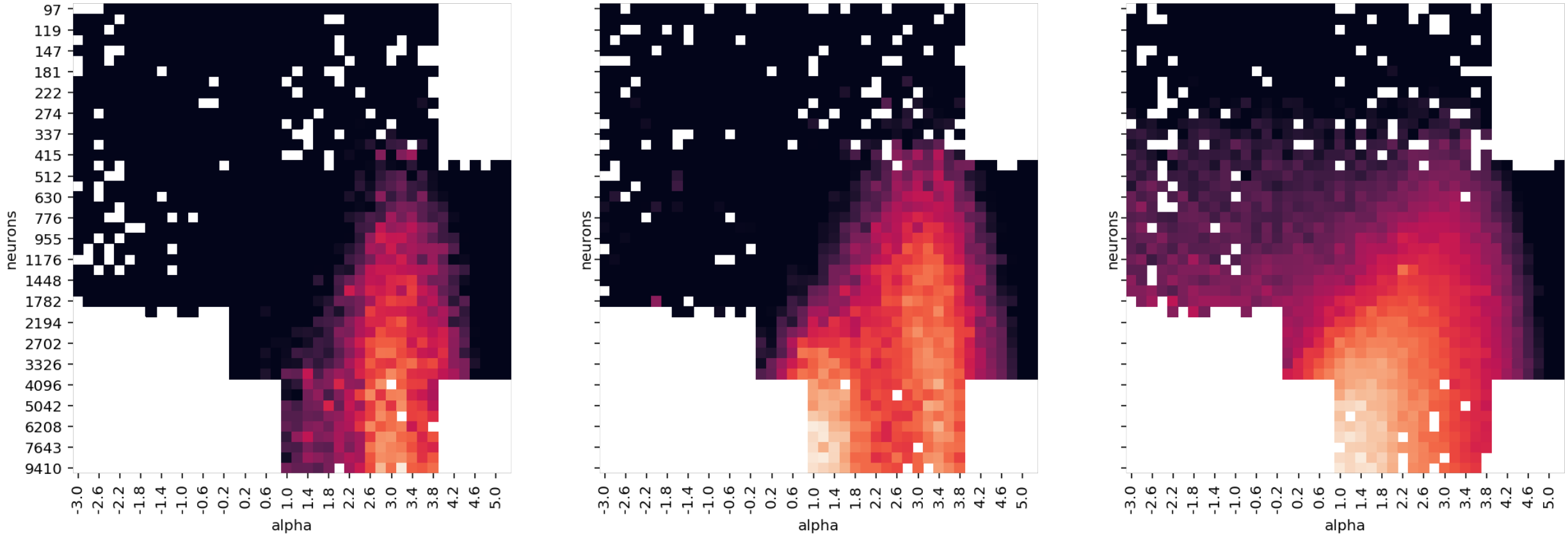
Our approach: Model pipeline

1. Local image patches are extracted from signature
2. Feature vectors are computed for image patches by a pre-trained general purpose DL network Inception21k
3. Extreme Learning Machine classifies feature vectors as coming from genuine/forged signature
4. Predictions are computed for feature vectors from image patches of validation signature
5. Signature prediction is majority vote over its local patch predictions

Our approach: Extreme Learning Machine (ELM)

- ELM is a single layer neural network with fixed input weights
- It has very fast training at huge model size, and easy to tune parameters
- Very efficient shallow learning method for summarising dataset

ELM Local signature features classification



ELM Local signature features classification

- Classification accuracy for 10%, 50%, 90% overlap
- Highest local feature validation accuracy is around 62%
- Larger overlap provides much larger dataset (more local feature per one signature image)
- With less data, optimal performance is at high regularization $\alpha=10^{3.0}$
- With more data, a new higher optimal performance is at lower $\alpha=10^{1.0}$
- Large training set of local feature provides natural model regularization, reducing the need for an external L2 regularization and improving model predictive performance

Results: Offline signature verification

- One user processed at a time, with 15 genuine signatures and 15 trained forgeries written by other users
- 30 total signatures split into 10 groups for 10-fold CV
- Results averaged over 10-fold CV for all users
- 90% overlap Equal Error Rate better than 2.45% from [12]

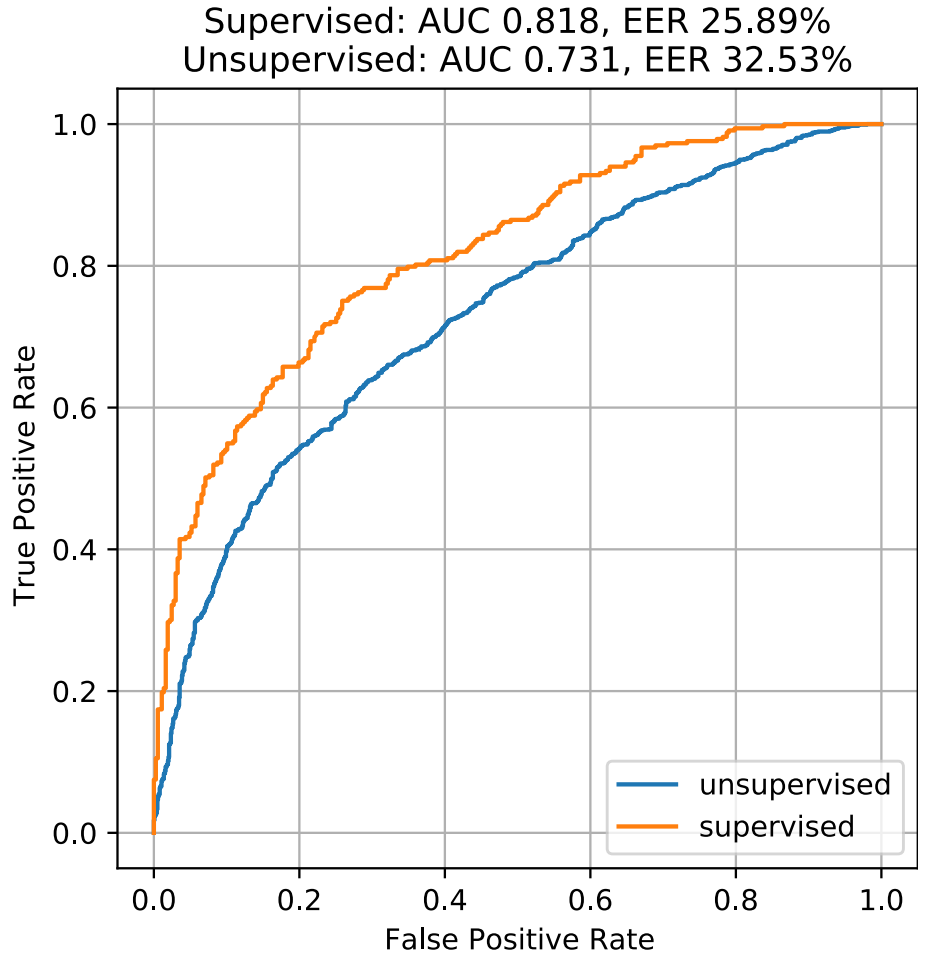
Overlap	10%	50%	90%
EER	7.11%	3.56%	2.31%

Results: General signature verification

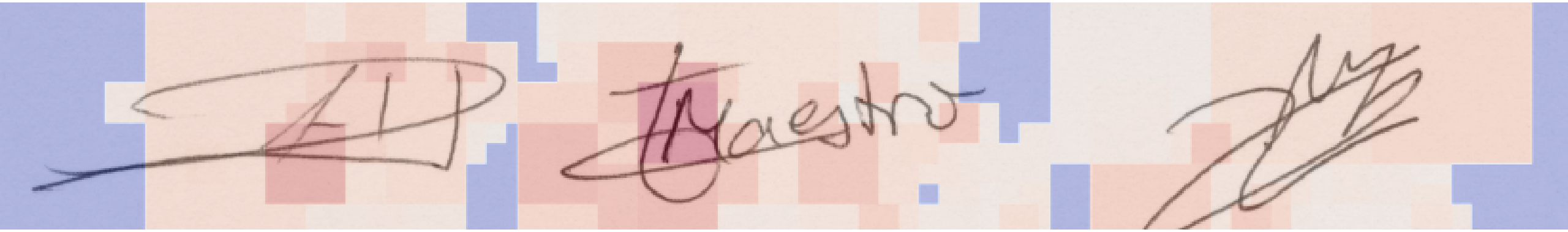
- Signatures of different users are pooled together with no distinction between users
- "Supervised" case splits data across signatures: some signatures of users go into training set, other signatures of same users go into validation set
- "Unsupervised" case splits data across users: all signatures of some users go into training, all signatures of other users go into validation
- "Unsupervised" case model need to detect forgeries in signatures of people it has never seen before!

Results: General signature verification

- Error rates are much higher at 25.9%, and 32.5% for unsupervised case
- Both error rates are above random guessing - it is possible to detect forgeries by just looking at a signature!
- More research needed, and a larger dataset: GPDSS10000



Thanks for your attention!



Feel free to send any questions to me (corresponding author)

anton.akusok@arcada.fi