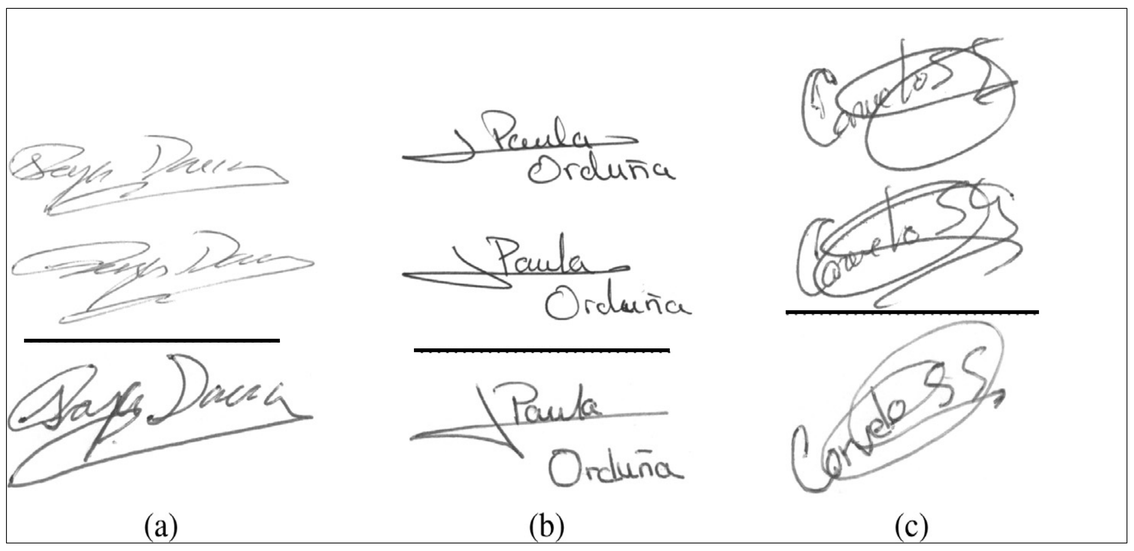
An Investigation of Feature Selection and Transfer Learning for Writer-Independent Offline Handwritten Signature Verification

Victor L. F. Souza[1], Adriano L. I. Oliveira[1], Rafael M. O. Cruz[2] and Robert Sabourin[2]

[1]Centro de Informática - Universidade Federal de Pernambuco and [2]École de Technologie Supérieure - Université du Québec

INTRODUCTION

HSV problem: verify whether a signature belongs to the claimed writer



BASIC CONCEPTS

Writer-Independent (WI) HSV systems:

In WI systems a single model is responsible for verifying signatures for all writers.

The verification process is carried out in the dissimilarity space resulting from the comparison between a questioned and a reference signature, through the Dichotomy Transformation

- Dichotomy transformation (DT): $\mathbf{u}(\mathbf{x}_q, \mathbf{x}_r) = \begin{bmatrix} |x_{q1} - x_{r1}| \\ |x_{q2} - x_{r2}| \\ \vdots \\ |x_{qn} - x_{rn}| \end{bmatrix}$

Allows to transform a multi-class problem into a 2-class problem.

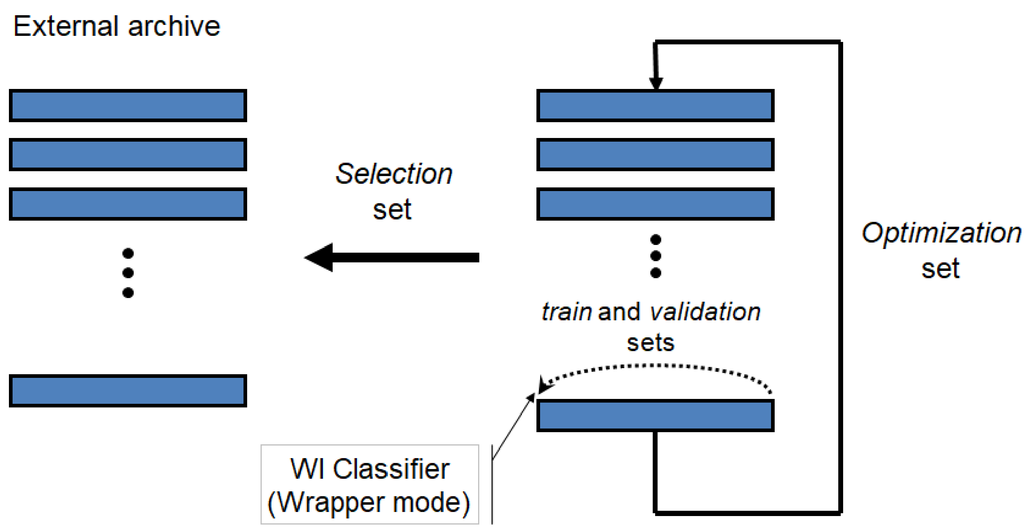
Feature selection using BPSO:

A feature selection technique based on binary particle swarm optimization (BPSO) can be used to obtain only the relevant dimensions on the dissimilarity space.

- **Fitness function**: The optimization is conducted based on the minimization of the Equal Error Rate (EER) of the WI classifier in a wrapper mode.

To decrease the chance of overfitting, a validation procedure can be used during the optimization process in order to select solutions with good generalization power.

- No validation (*NV*)
- Partial validation strategy (*PV*)
- Global validation strategy (*GV*)
  - This can be accomplished by storing the best validated solutions in an external archive



EXPERIMENTS

Objectives:

- Investigate both the use of feature selection through BPSO and the effectiveness of the overfitting control strategy.
- Analyze whether the space generated by the feature selection can be used in a transfer learning context.

Proposed method:

COMPARISON OF EER CONSIDERING THE PRESENTED MODELS, IN THE GPDS-300 DATASET (ERRORS AND STANDARD DEVIATIONS IN %)

Approach	#features	EER
No feature selection	2048	3.47 (0.15)
$BPSO_{NV}$	1124	3.76 (0.07)
$BPSO_{PV}$	1120	3.64 (0.08)
$BPSO_{GV}$	1140	**3.46 (0.08)**

Transfer Learning:

COMPARISON OF EER CONSIDERING THE PRESENTED MODELS, IN A TRANSFER LEARNING CONXTEXT IN THE CEDAR AND MCYT DATASETS (ERRORS AND STANDARD DEVIATIONS IN %)

Approach	#features	EER_{CEDAR}	EER_{MCYT}
No feature selection	2048	3.32 (0.22)	2.89 (0.13)
$BPSO_{NV}$	1124	4.00 (0.17)	2.69 (0.13)
$BPSO_{PV}$	1120	3.98 (0.25)	2.56 (0.05)
$BPSO_{GV}$	1140	**3.27 (0.22)**	**2.48 (0.23)**

Comparison with the state of the art in the MCYT dataset:

COMPARISON OF EER WITH THE STATE OF THE ART IN THE MCYT DATASET (ERRORS AND STANDARD DEVIATIONS IN %)

Type	HSV Approach	#Ref	#Models	EER
WD	Vargas et al. [38]	10	75	7.08
WD	Ooi et al. [39]	10	75	9.87
WD	Zois et al. [21]	5	75	6.02
WD	Hafemann et al. [4]	10	75	2.87 (0.42)
WD	Zois et al. [30]	5	75	3.97
WD	Hafemann et al. (fine-tuned) [22]	10	75	3.40 (1.08)
WD	Okawa [31]	10	75	6.40
WD	Zois et al. [33]	5	75	3.52
WD	Zois et al. [24]	10	75	1.37
WI	Zois et al. [27]	5	1	3.50
WI	Souza et al. [17]	10	1	2.89 (0.13)
WI	$BPSO_{GV}$	10	1	2.48 (0.23)

CONCLUSIONS

Experimental results showed that not using a validation stage is worse than using the partial validation, which in turn is worse than using the global validation strategy.

Another aspect that can be observed is the redundancy of the features in the dissimilarity space, since the models with a validation stage managed to obtain a better EER using only almost 55% of the total number of features.

The space generated after feature selection can actually be used in a transfer learning context, which means, the optimized feature representation can generalize better across different datasets

ACKNOWLEDGMENTS

This work was supported by CNPq, FACEPE and ÉTS Montréal.

25th International Conference on Pattern Recognition
10 - 15 July 2021 Milan Italy