Test-Time Augmentation for Document Image Binarization

Adrian Rosello, Francisco J. Castellanos, Juan P. Martinez-Esteso, Antonio Javier Gallego and Jorge Calvo-Zaragoza

Gothe might

1. Introduction

This research probes the efficacy of Test-Time Augmentation (TTA) for improving document binarization in **data-scarce scenarios** within the realm of digital transcription.

The study highlights TTA's potential to enhance performance, especially in datalimited situations, but indicates the need for further refinement of TTA strategies for complex tasks.

2. Methodology

3. Experimental Setup



The image to be binarized is disturbed following data augmentation strategies. Each augmented version is independently binarized and then the results are aggregated to give a more robust result.



	Vote	Counts the number of cases in which a pixel was predicted as ink, then final prediction is based on the percentage of ink votations						
y		vote50 (50%)	vote75 (75%)	votel00 (100%)				
60	Average	Computes the	average confic	lence hetween	all the			

Concerning the data augmentation techniques utilized for training the augmented models as well as the ones applied for TTA, they were randomly applied in the form that several transformations could be used to the selected patch.





confidences, then a threshold is applied to determine the class.

Maximum Takes the maximum confidence value for each pixel and then, the resulting probabilistic map is processed by a threshold.

4. Results



5. Conclusions

Our experiments over 5 corpora of document collections reveal a **modest improvement** by using TTA for binarization.

The improvement is greater as the amount of annotated data is reduced, thereby making this technique an interesting solution to be <u>further</u> <u>studied</u>.

Scenario	Best combination policy				
Limited Data	max-policy & vote50				
Sufficient Data	votel00				

6. Future Work

Base case $(\mathbf{F}_1 \ \%)$										
Non-augmented model		18.2	22.7	26.8	25.8	28.5	38.7	56.9	61.2	33.3
Augmented model	24.7	14.7	23.3	20.7	32.1	31.8	50.6	57.0	58.3	34.8
Improvement rate of TTA (%)										
Non-augmented model	2.7	3.3	4.5	3.4	1.1	1.5	2.8	0.3	0.2	1.7
Augmented model	5.3	5.6	3.1	2.3	1.9	1.1	-0.2	0.0	2.7	1.8



- Combination of TTA with meta-learning techniques
- Few-shot-based architectures blended with TTA to improve binarization quality
- Explore further augmentation transformations

7. Acknowledgement

This work was supported by the I+D+i project TED2021-132103A-I00 (DOREMI), funded by MCIN/AEI/10.13039/501100011033.









