

# dhSegment

A generic deep-learning approach  
for document segmentation

**Sofia Ares Oliveira**

Benoit Seguin, Frederic Kaplan

Ecole Polytechnique Fédérale de Lausanne, Switzerland



CIE  
Gen-  
tesimi

N. della Mappa	N. della subalter.	POSSESSORI	Denominazione dei Pezzi di terra	QUALITA'	SUPERFICIE	
					Classi	Pertiche Centesime
			<i>Rovaglia f. degli.</i>			
75.	1.	<i>Primi Pietro</i>	3275.	<i>Bottega unia d'Affitto</i>		
	2.	<i>Secunda Stefano</i>	<i>e. s. p.</i>	<i>Superiore d'Affitto.</i>		
76.	1.	<i>Primo suddetto</i>	3276.	<i>Casa d'Affitto</i>		
	2.	<i>Secondi ecclesiastici</i>	<i>e. s. p.</i>	<i>Casa d'Affitto superiore</i>		
77.		<i>Secondi suddetto</i>	3277.			
78.		<i>Padreuzza Giuseppe e Vittore</i>	3278, 3279.	<i>Casa e Bottega d'Affitto</i>		
79.		<i>e Nova Margherita</i>	3282.	<i>Casa di propria abitazione</i>		
80.	1.	<i>Primi Pietro</i>	3275.	<i>Casa in due stanzette per</i>		
	2.	<i>Secunda Stefano</i>	3276.	<i>propria d'Affitto</i>		
				<i>propria superiore d'Affitto</i>		
81.		<i>Secunda Stefano</i>	3288, 3284.	<i>Casa di propria abitazione</i>		
82.		<i>Franco Antonio g. Giuseppe</i>	3285.	<i>Casa d'Affitto ad uso di</i>		
				<i>bottega.</i>		
83.	1.	<i>Paganò da Belluno</i>	3286.	<i>Casa d'Affitto</i>		
	2.	<i>Farina Antonio g. Giuseppe</i>	3285.	<i>propria superiore d'Affitto.</i>		
84.	1.	<i>Secondi Giuseppe</i>	3287.	<i>Casa d'Affitto a propria abitazione</i>		
	2.	<i>Paganò da Belluno</i>	3288.	<i>Casa superiore d'Affitto.</i>		
85.		<i>Barada Francesco</i>	3309, 3310.	<i>Casa d'Affitto.</i>		
86.		<i>Giuseppe Feliciano e Bernardi Luigi</i>	3311.	<i>Casa d'Affitto.</i>		
		<i>indivisi.</i>				
87.	1.	<i>Giuseppe Feliciano</i>	3312.	<i>Casa d'Affitto</i>		
	2.	<i>Giuseppe suddetto e Bernardi Luigi</i>	3311.	<i>propria superiore d'Affitto.</i>		
88.		<i>Faccarini Caterina Vedova del</i>	3313, 3314, 3315.	<i>Casa d'Affitto.</i>		
		<i>g. Giovanni Formani.</i>				

**D**omine ihesu xpe. **28** *sextani.*  
**D**eus in adiutorium  
 meum intende.  
**G**loria p<sup>ri</sup>. **ymnus.**  
**H**ora sexta ihesus est crucis  
 clauatus et est cum latro  
 iudicatus et est cum latro  
 iudicatus. Pre tor  
 mentis sciens felle satucatus  
 acutus crumen diluit sic iudifi  
 catus. **o** adoramus te. **R.** Quia. **oro.**  
**D**omine ihu xpe. **ad nonan.  
**D**eus in adiutorium me  
 um intende.  
**D**omine ad adiuvandum me  
 festina.  
**G**loria p<sup>ri</sup>.**



*En prioune come a cestuy cy*



et ta Baillance / se tay tousiours trouue preudhomme & loyal  
 Tu es hōme dauctorite et bien cōgneu en ces marches. Je tē  
 constitue roy / a te donne tout le pays / a ne te baille aultre char  
 ge q̄ de aymer Vertu & tousiours suyuir hōneur. Quāt Hispan  
 ouyt le don q̄ Hercules tui faisoit il cheut a ses piedz & le remer  
 cia. Puis se excusa de si grāt honneur. Mais Hercules tui dist  
 quil vouloit quil fust ain si / a tui bailia certain nombre de ses  
 gens pour se servir / puis se fist partir de sa cōpaigie a grāds  
 regretz. Et Hispan sey alla par to<sup>s</sup> les pays q̄ Hercules auoit  
 isse cōquis sur Berion & sur Lacus. Et des lors en auant se  
 pays fut nomme espaigne de son nom. Dōt me fait ay a tant  
 de ceste conqueste despaigne & diendray a pat ser des armes &  
 Baillances q̄ Hercules fift en lōbardie / a de la mort de Lac.  
 Comment Hercules se combatit contre les geantz de cre  
 mone & les vainquit.



Il sont maintenāt les emperours / les roys / les  
 ducs / les soudans q̄ se pourront a cōpartager de  
 Vertueuse liberalite sebiabte a celle de Hercu  
 les. Les hōmes au iourd'uy cōbatēt tui cōtre  
 l'autre & font assez de conquestes / mais tout attri  
 buēt a leur sigulier puffit. Ilz ne ressembēt pas a Hercules q̄  
 ne cōbatoit q̄ pour le bien cōmun du mōde. D noble  
 Herc. R iii

ERYCI PVTEANI  
 ELOQVENTIÆ  
 AVSPICIA  
 SECVNDA.  
 ORATIO  
 AD MEDIOLANENSES,

**L** ANVM Antiquitas, Auditores, Bicipitem, velut Prudentiæ quamdam imaginem, typumque finxit; vt in rebus consiliisq; non tantum quid actum, sed quid agendum quoque esset, pari lance expensum haberemus; & in futurum præmittere festinantes cogitationes solliciti, & ad stricto quodam consilii freno cohibere ac reflectere cati, ad ea quæ præterita sunt. Rectè. Ratio enim transacti forma quædam futuri est; vt si quid feliciter successerit, hoc ipsum securo exemplo affectare; si quid præter animi sententiam passimur, cuitare studeamus. Ego verò cum in hoc studiorum principio, & non lani dicam, sed Suadæ limine Orationem more maiorum habiturus essem, Muneris publici causâ; æratem hanc meam, dicendique verecundiam prudentiæ quamdam affectatione supplere conatus sum; &, Bifiõtem cogitans, elapsi successum anni breuiter referre, & instantis simul rationem disputare constitui; vt, bono omine, vos quoque, Auditores

S. A. Oliveira, B. Seguin, F. Kaplan, dhSegmKnt - JCFHR 2018

**Comment limiter les pannes des satellites**  
**Allers et retours sur l'affaire des camps**  
 Rumeurs et démentis sur les camps de concentration en Bosnie traduisent plus de peur que d'embarras. Analyse.

**ESPACE**  
 Les problèmes du satellite Eureka, dimanche dernier par Atlantide, résultent vraisemblablement d'erreurs dans les programmes informatiques. Le Centre de Darmstadt parviendra probablement à les corriger. En revanche, il restait impuissant face à des problèmes mécaniques ou électriques. Pour lui, les engins en orbite vivent dans un environnement très hostile. Les ingénieurs doivent recourir à des technologies particulières pour concevoir sur nombreux mécanismes impossibles, par exemple, d'utiliser un lubrifiant comme de l'huile, qui s'évaporerait ou se figerait immédiatement. Pour limiter les risques d'une panne inévitable, les satellites subissent de nombreux tests, notamment dans un « simulateur spatial ». L'une de ces machines se trouve au Centre technique de l'Agence spatiale européenne, en Hollande. Elle génère le vide, le froid et le plus grand rayon de soleil artificiel du monde... (red.)

**EX-YOUGOSLAVIE** L'Occident face à la guerre  
 Après une semaine de rumeurs persistantes sur l'existence de camps de déportation en ex-Yugoslavie dans lesquels à l'abri des regards se dérouleraient les pires horreurs, les gouvernements et les institutions humanitaires tendent à relativiser la véracité de ces allégations. Le Département d'Etat américain avait dans un premier temps semblé confirmer les rapports de presse sur des camps de concentration pour détenus musulmans aux mains des milices serbes en Bosnie. Il a fait hier marche arrière dans ses déclarations. Le CICR, prudent par fonction, estime que certains camps échappent à ses investigations et que prochainement pour cette raison il ne sait pas ce qui s'y commet. L'émotion dans la presse et dans les opinions publiques est cependant plus grande qu'elle n'a jamais été depuis le début des guerres yougoslaves. Elle parvient presque à troubler la sérénité vacancière et la délectable frivolité des exploits olympiques. Les témoignages



Faut-il un langage - une action - plus ferme pour arrêter cette mortelle dérive? (Keystone)

**ESTIVALES**  
**Montbard, cité de Buffon**  
 Sur les traces du grand naturaliste, Intendant du Jardin du Roi, fidèle à sa ville natale, en Bourgogne.  
 Lire en page 2

**ÉTRANGER**  
 Des dizaines de milliers de Noirs défilent derrière Nelson Mandela 3

**ÉCONOMIE**  
 L'Allemagne a franchi la barre des trois millions de chômeurs 5

**SPORTS**  
 Saut en longueur: Lewis/Powell, un duel historique 11

**SUISSE**  
 Vellerau va fêter le dixième anniversaire de sa « déclaration d'indépendance » 13

**VAUD**  
 Montreux: coup de frein sur les congrès 16

**GENÈVE**  
 Chamboulements au marché aux puces 17

Décès  
 Carnet  
 Cinéma  
 Radio/TV 20  
 21  
 22  
 23

**EDITORIAL**  
**Villes interdites**  
 par Pascal Garcin  
 «Ville privée: entrée interdite.» Le panneau de signalisation dûment normalisé par les accords internationaux n'existe pas encore. Mais il pourrait voir le jour prochainement si s'affirmait une tendance observée notamment aux États-Unis. Il s'agit de la création d'associations municipales privées. Les adhérents s'acquittent de taxes et soustraient à des règlements. En échange, ils disposent d'infrastructures (routes, équipements collectifs, transports) et de services (police privée, télévision, etc.) Ils réalisent donc sur une base purement privée et contractuelle ce que l'Etat, et plus particulièrement ceux qui ont le devoir d'offrir dans le cadre de la société de services municipaux.

**Certains Américains veulent privatiser les villes. Une façon de créer une société à deux vitesses.**  
 Les villes interdites, en effet, ceux qui ont des besoins particuliers en matière de sécurité ou de tranquillité. L'assurance n'existe pas eux-mêmes les coûts imposés par la satisfaction de ces besoins? C'est ce que font par exemple tous ceux qui placent leurs enfants dans des écoles privées, ou qui recourent à des services privés de protection. Les villes privées, par le caractère global des services qu'elles proposent présentent toutefois la caractéristique particulière d'être un embryon de société concurrente des sociétés publiques. Autant dire que dans ce cadre concurrentiel, ceux qui ont les moyens peuvent bénéficier de services plus efficaces, les autres doivent se contenter de ce qui est offert par l'Etat. Les choses pourraient en rester là si le processus ne présentait pas un effet pervers. Car à mesure que s'amplifie le succès de ces nouvelles villes, s'agrandit le cercle de ceux qui en deviennent citoyens, et qui militent tout naturellement en faveur d'une diminution des dépenses publiques. Le risque de délabrement des sociétés publiques qui en résulte constitue alors une véritable incitation à créer de nouvelles villes privées. Au sens propre, c'est une société à deux vitesses qui serait ainsi créée. Par nature, ce type de société ne peut être qu'instable, et donne lieu à des explosions de violence d'autant plus brutales et fréquentes qu'est importante la différence de vitesse entre les deux sociétés. Or ces violences, elles aussi, présentent un coût qu'il faut mettre dans la balance lorsque l'on analyse au niveau de la société entière les avantages et les inconvénients du secteur public et du secteur privé.

**JAPON** Chute de la bourse  
**Les banques sont dans une zone dangereuse**  
 Pour les sauver, l'Etat pourrait entrer dans leur capital.  
 A Tokyo, les réserves latentes sur les titres que détiennent les banques japonaises dans leur portefeuille ne permettraient déjà plus de couvrir complètement les prêts douteux qu'elles détiendraient ailleurs, selon une étude de Borden & Cie. Si le Nikkeï continuait sa chute, le pire serait ainsi à craindre. Toutefois, estime Borden & Cie, il est probable que le gouvernement interviendrait alors pour empêcher une catastrophe financière, en entrant d'une manière ou d'une autre dans le bilan des banques.

**SUISSE** Votation fédérale  
**Les défenseurs des petits paysans prennent l'offensive**  
 Combattue par un référendum, la révision du droit foncier rural visait aussi le monde paysan.  
 Hier à Berne, ce sont les défenseurs du nouveau droit foncier rural qui ont ouvert la campagne. Ils posent pour principe de base que les terres cultivées doivent demeurer aux mains de ceux qui les exploitent. Ils plaident, par conséquent, en faveur des mesures, prévues par cette révision, qui visent à empêcher la spéculation sur ce qu'ils qualifient de « terre morte » à savoir: la terre cultivable. (Réd.)  
 Lire l'article de Jean-Marie Crevatier en page 13



BERLINO

MARATTA, Carlo.

KUPFERSTICHKABINETT  
Staatliche Museen  
n. 21077

Studio per divinità fluviali.  
195 x 268

Corpus Gernsheim

52101





# Motivation

The **diversity** of historical document processing tasks prohibits to solve them one at a time

→ Design a **generic approach** in order to handle this variability

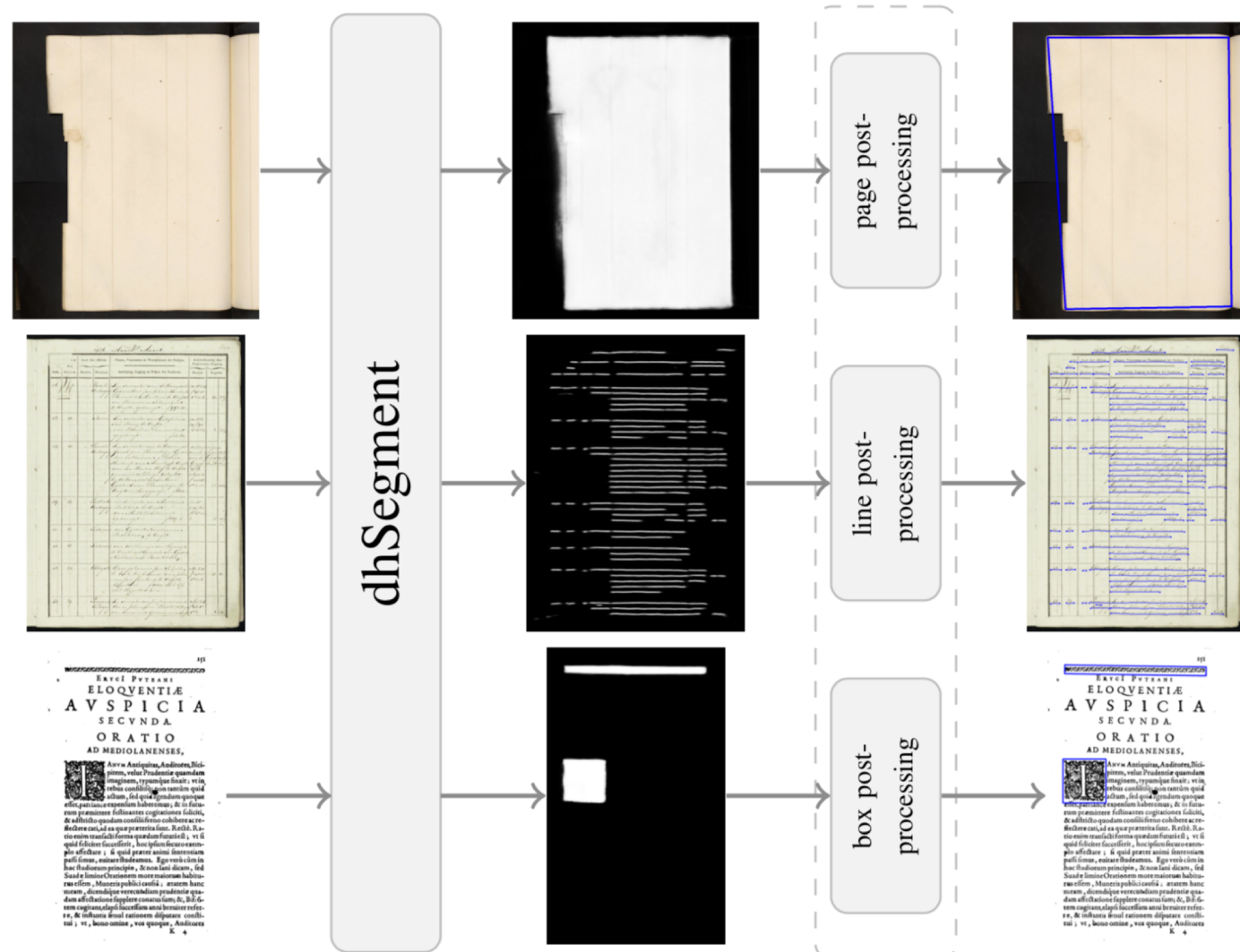
Growing interest in digital humanities research

→ Increasing need for **simple** to use, **flexible** and **efficient** tools to process and analyse historical documents, usable by non-specialists



# dhSegment

# System overview

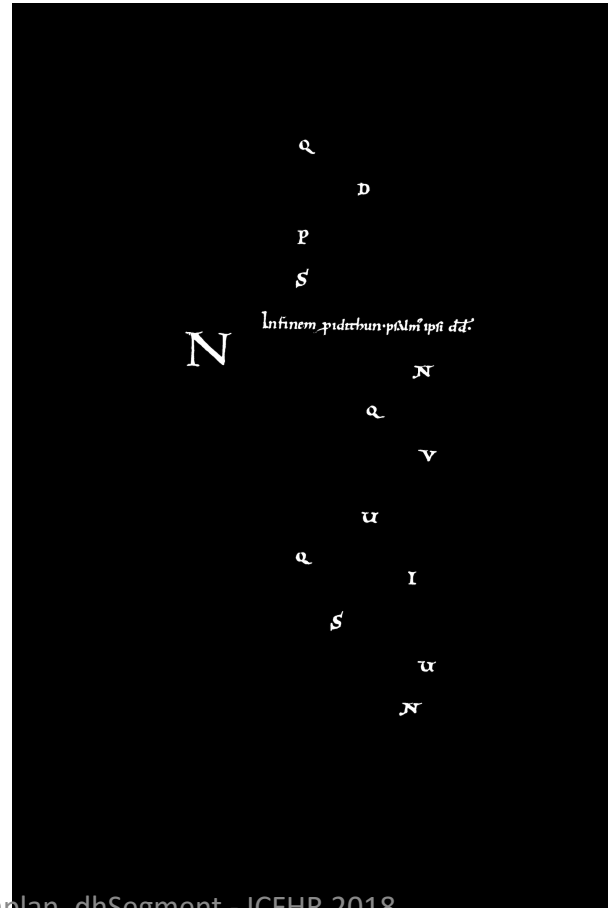
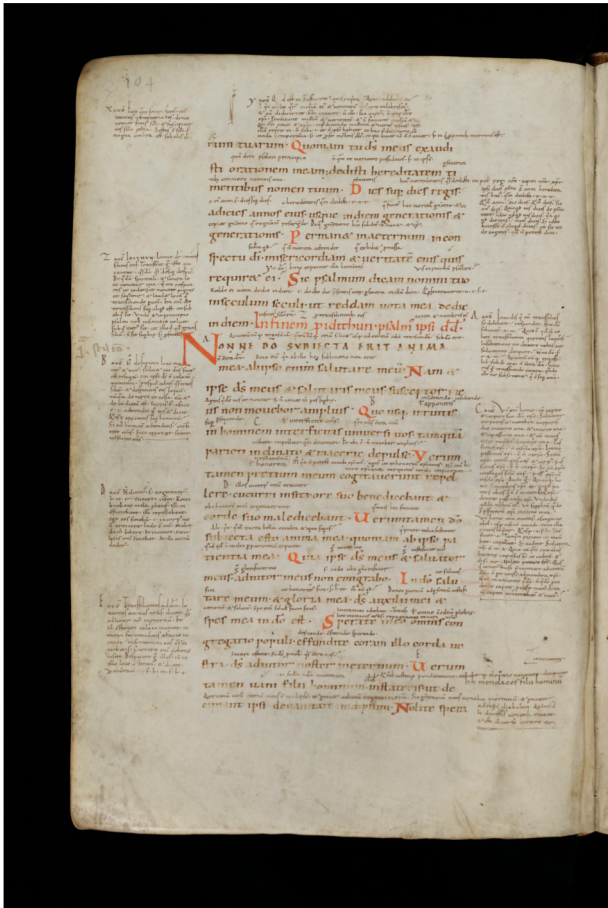






# Training

Create a **binary mask** for each class



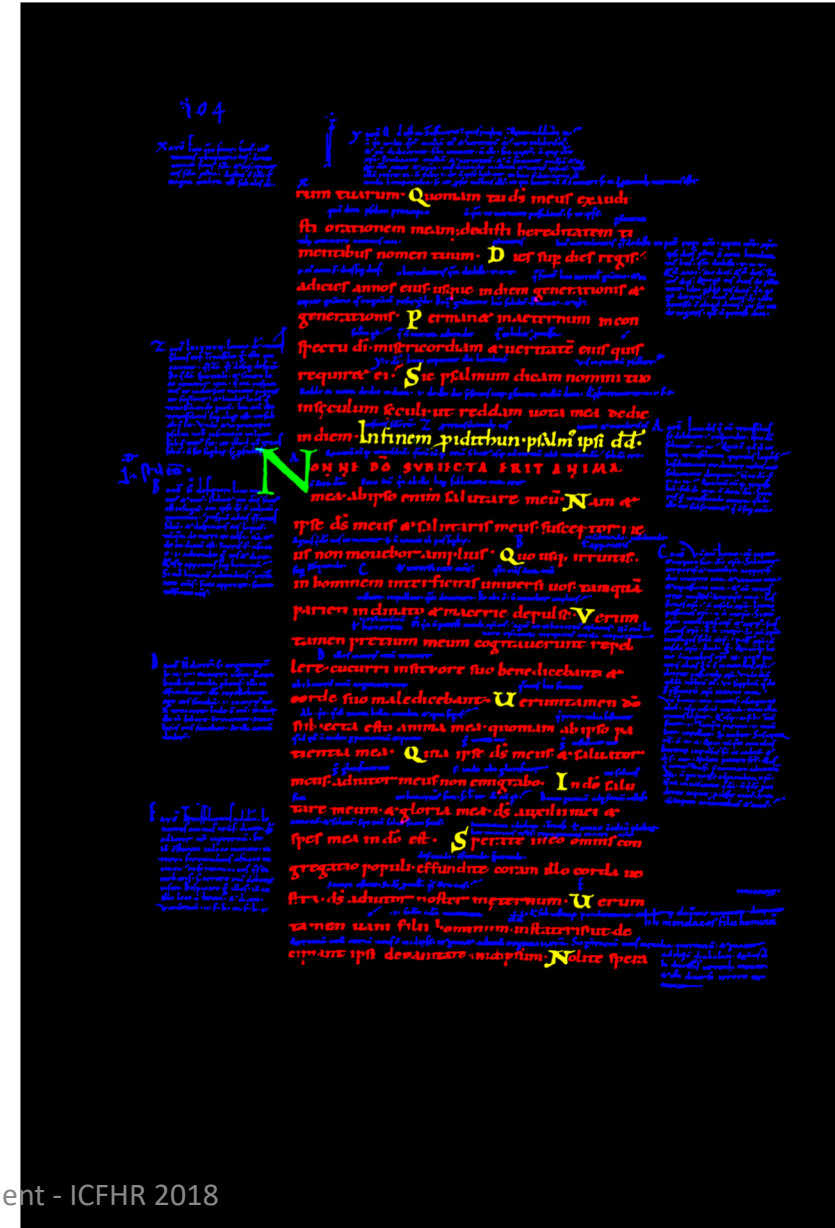
Images from DIVA-HisDB dataset



# Training

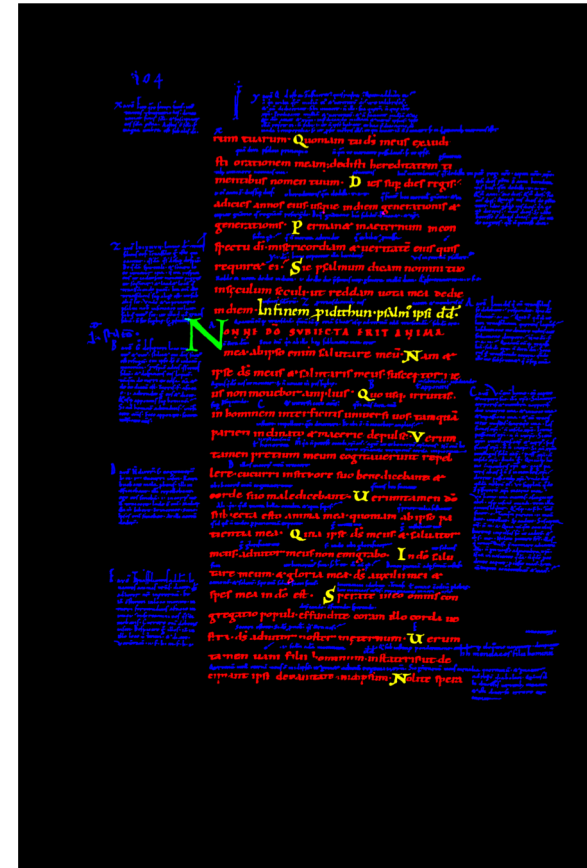
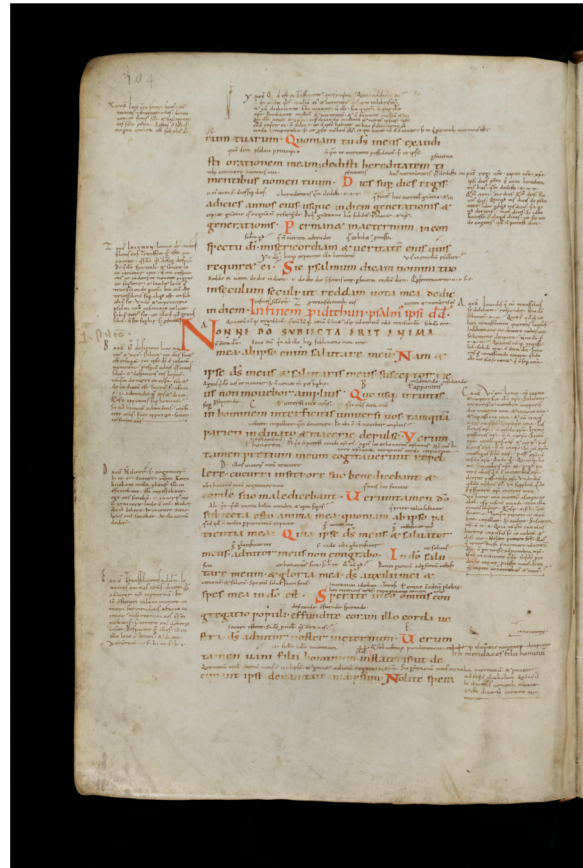
Combine the binary masks to make a labelled image

Each pixel is assigned one or several labels



# Training

Input original image + labelled image to the network

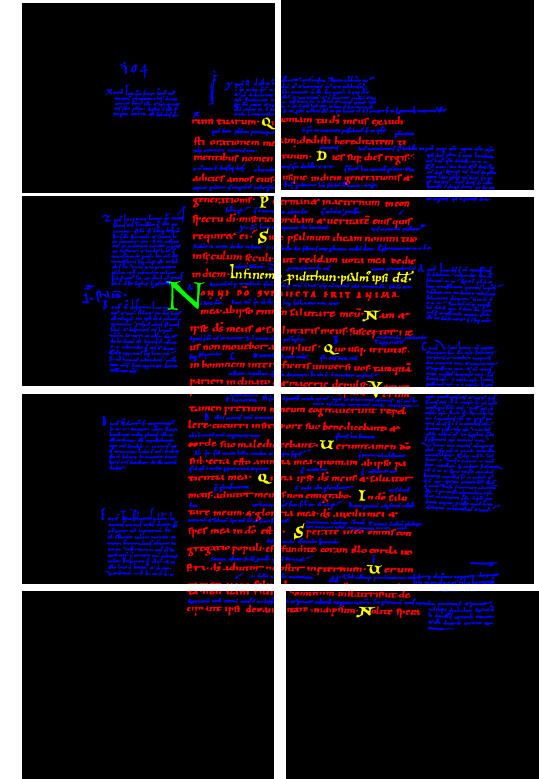




# Training

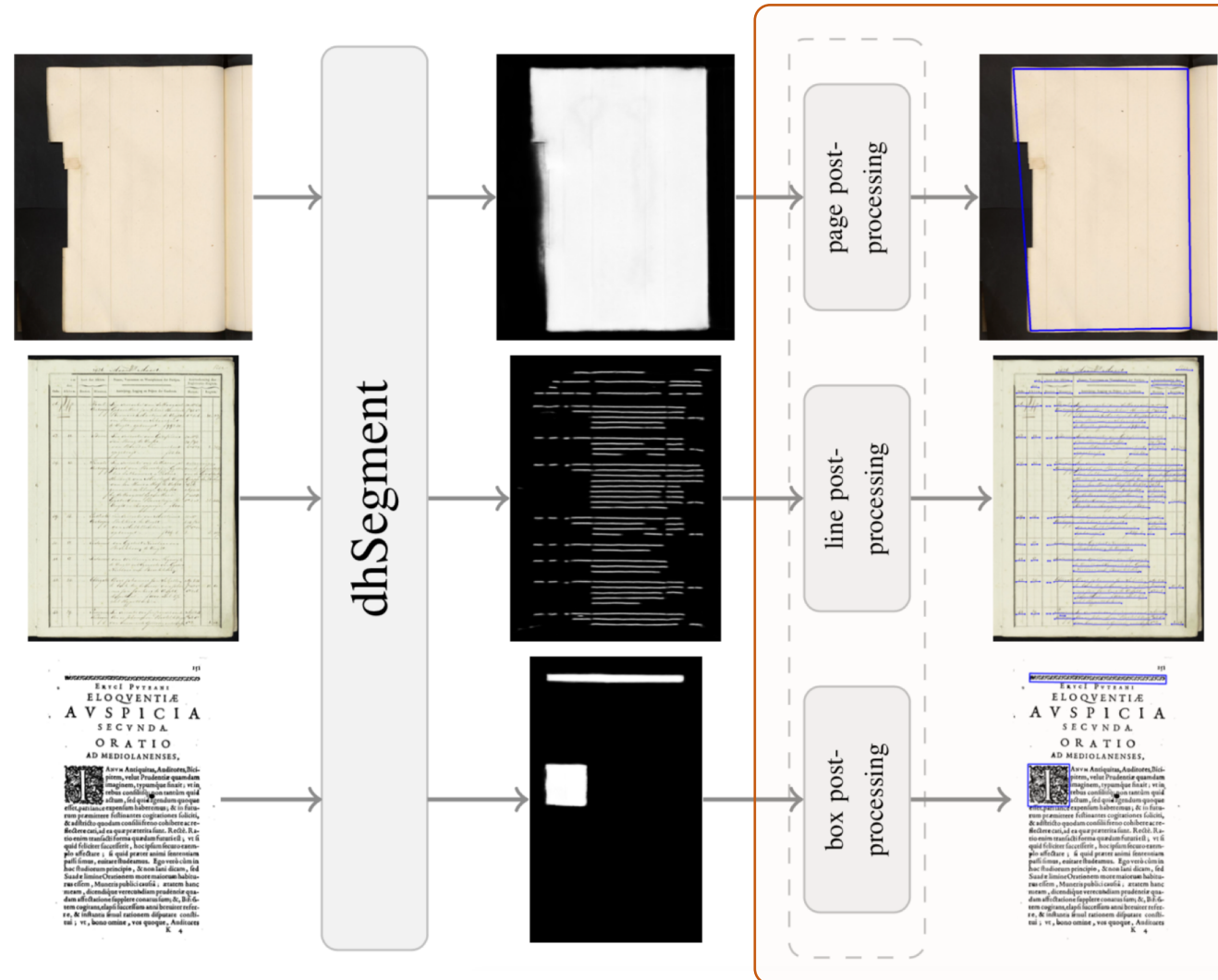
Patches can be used in order to optimize batch training

Data augmentation on the fly (rotation, scaling, mirroring, ...)



# Object extraction / detection





# Object extraction / detection

dhSegment outputs **per-pixel probabilities**

Object **extraction is done in the post-processing** step, which is independent of the neural network

The provided post-processing blocks (in the experiments) are **limited to simple and standard operations**

thresholding, morphological operations, connected components analysis and shape vectorization



# Experiments

# Experiments

Page extraction (READ-BAD)

Baseline detection (cBAD, ICDAR2017)

Document layout analysis (DIVA-HisDB, ICDAR2017)

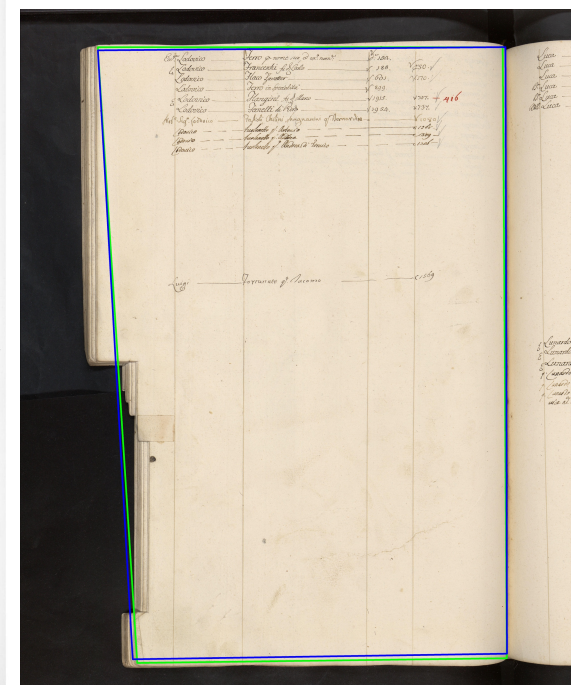
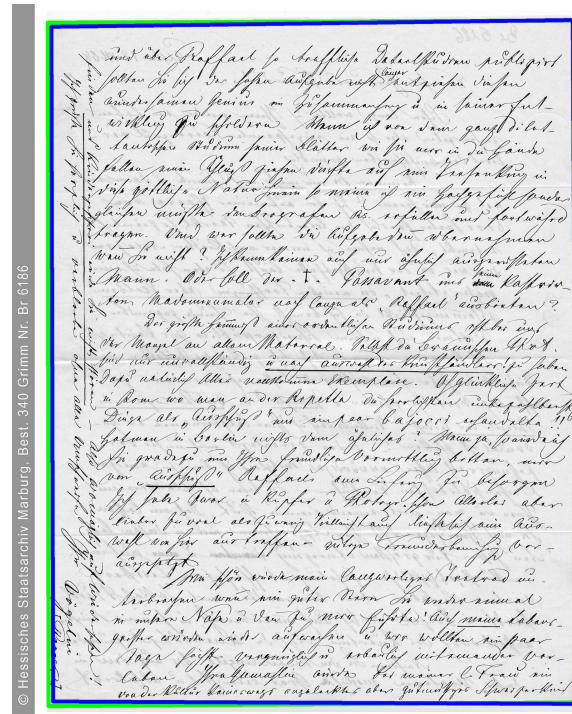
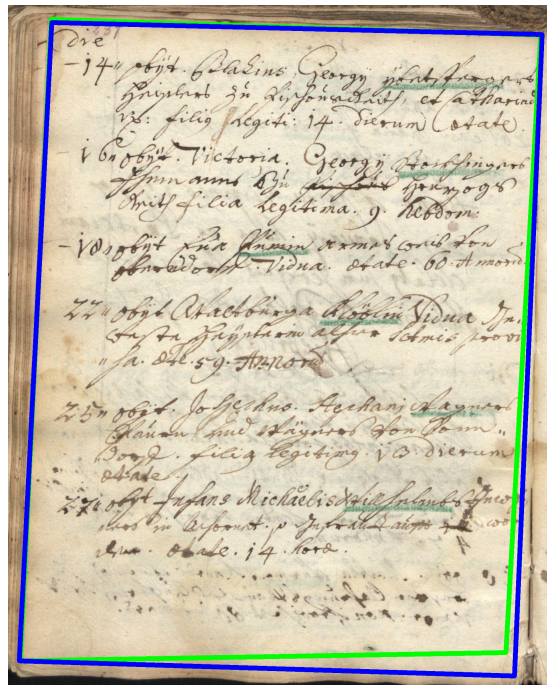
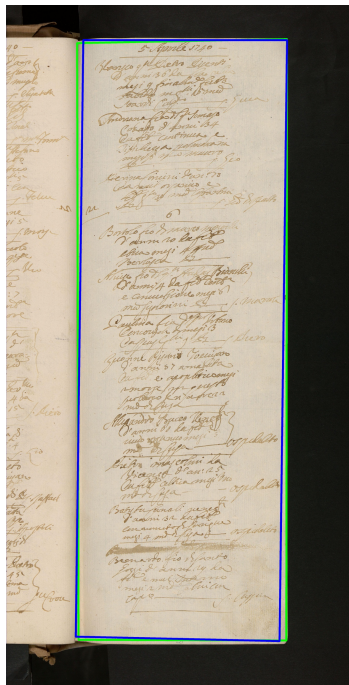
Ornament extraction (private)

Photograph extraction (private)



# Page extraction

Dataset : sample from READ-BAD [1] annotated by [2]



[1] T. Grüning et al. "READ-BAD: A new dataset and evaluation scheme for baseline detection in archival documents," 2018

[2] C. Tensmeyer et al. "Pagenet: Page boundary extraction in historical handwritten documents," 2017

# Page extraction

RESULTS FOR THE PAGE EXTRACTION TASK (mIoU)

Method	cBAD-Train	cBAD-Val	cBAD-Test
Human Agreement	-	0.978	0.983
Full Image	0.823	0.831	0.839
Mean Quad	0.833	0.891	0.894
GrabCut	0.900	0.906	0.916
PageNet [1]	0.971	0.968	0.974
<b>dhSegment (quads)</b>	<b><math>0.98 \pm 6e^{-4}</math></b>	<b><math>0.98 \pm 8e^{-4}</math></b>	<b><math>0.98 \pm 8e^{-4}</math></b>

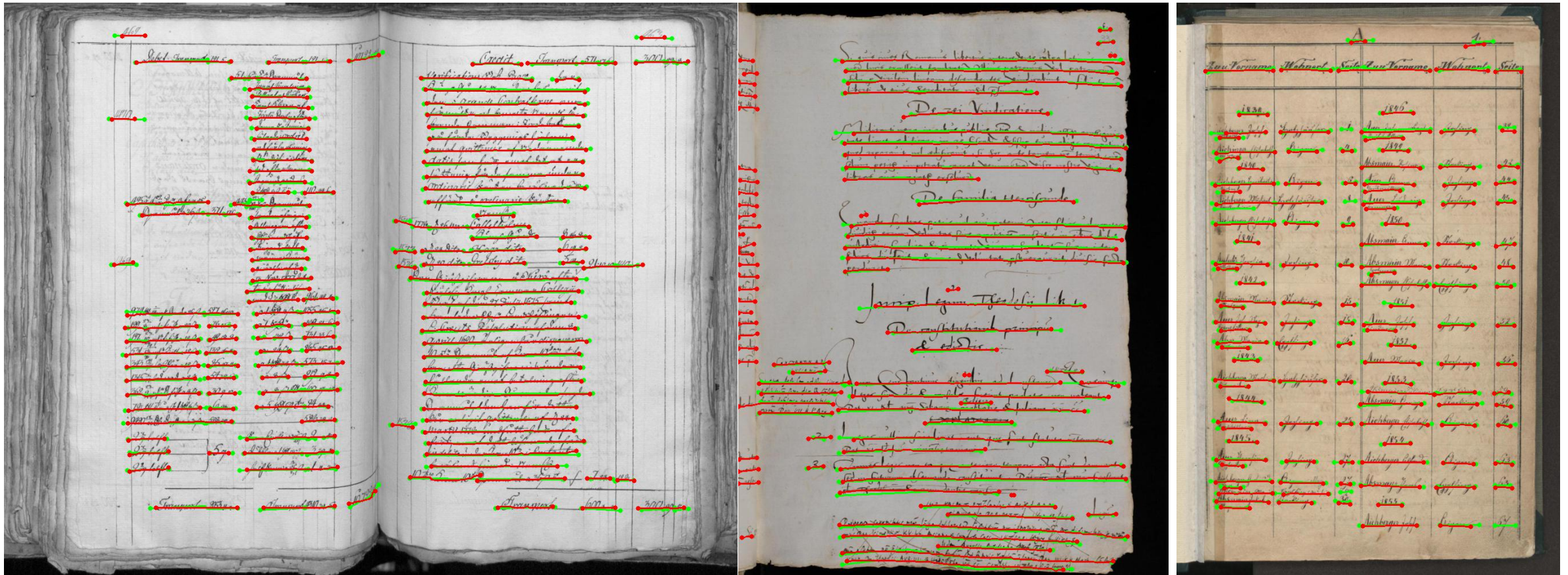
Train : 1635 images

Test : 200 images



# Baseline detection

Dataset : READ-BAD



T. Grüning et al. "READ-BAD: A new dataset and evaluation scheme for baseline detection in archival documents," 2018

S. A. Oweira, B. Segura, T. Kaplan, unsegmented, ICFHR 2018



# Baseline detection

RESULTS FOR THE cBAD : ICDAR2017 COMPETITION ON BASELINE  
DETECTION [1] (TEST SET)

Method	Simple Track			Complex Track		
	P-val	R-val	F-val	P-val	R-val	F-val
LITIS	0.780	0.836	0.807	-	-	-
IRISA	0.883	0.877	0.880	0.692	0.772	0.730
UPVLC	0.937	0.855	0.894	0.833	0.606	0.702
BYU	0.878	0.907	0.892	0.773	0.820	0.796
DMRZ	<b>0.973</b>	<b>0.970</b>	<b>0.971</b>	<b>0.854</b>	0.863	<b>0.859</b>
<b>dhSeg.</b>	0.88±.023	<b>0.97±.003</b>	0.92±.011	0.79±.021	<b>0.95±.005</b>	<b>0.86±.011</b>

\* Current state-of-the-art achieves 0.978 F-val and 0.922 F-val for simple and complex track respectively [2]

Simple track

Train : 216 images

Test : 539 images

Complex track

Train : 270 images

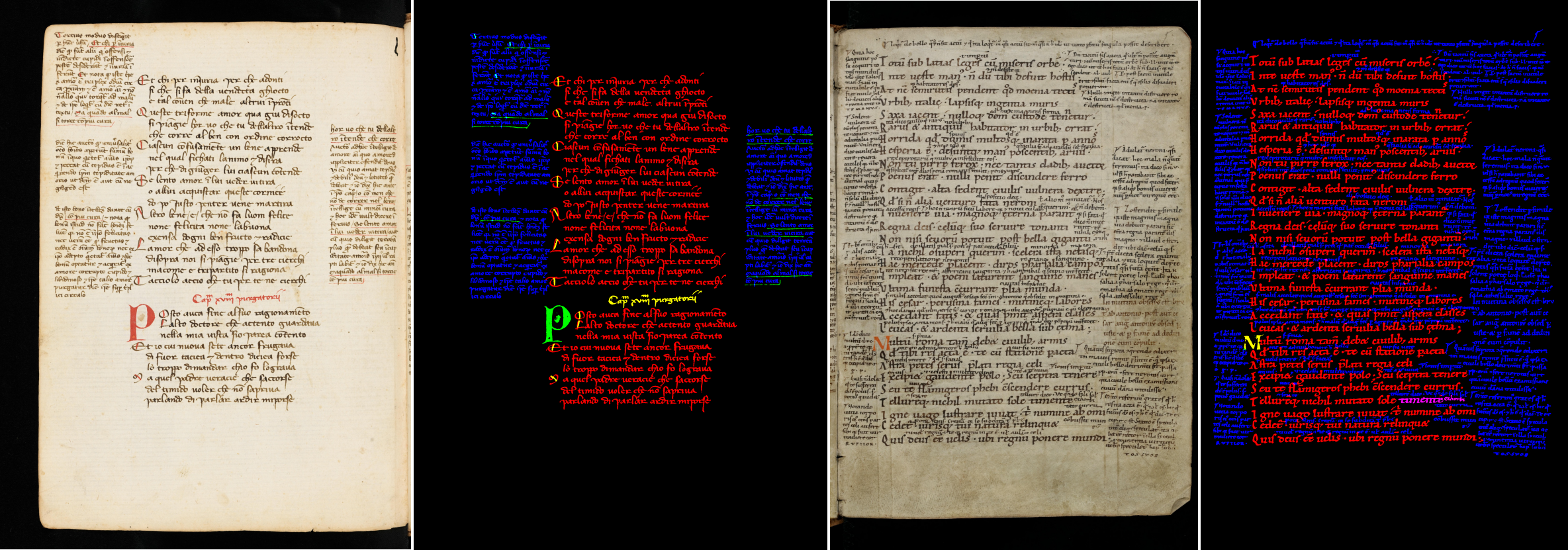
Test : 1010 images

[1] M. Diem, et al. "cBAD: ICDAR2017 Competition on Baseline Detection," 2017

[2] T. Grüning, et al. "A Two-Stage Method for Text Line Detection in Historical Documents," 2018

# Document layout analysis

## Dataset : DIVA-HisDB



F. Simistira, et al. "DIVA-HisDB: A precisely annotated large dataset of challenging medieval manuscripts," 2016. 22

# Document layout analysis (ICDAR2017)

RESULTS FOR THE ICDAR2017 COMPETITION ON LAYOUT ANALYSIS  
FOR CHALLENGING MEDIEVAL MANUSCRIPTS [1] - TASK-1 (IoU)

Method	CB55	CSG18	CSG863	Overall
System-1 (KFUPM)	.7150	.6469	.5988	.6535
System-6 (IAIS)	.7178	.7496	.7546	.7407
System-4.2 (MindGarage-2)	.9366	.8837	.8670	.8958
System-2 (BYU)	.9639	.8772	.8642	.9018
System-3 (Demokritos)	.9675	.9069	.8936	.9227
<b>dhSegment</b>	.974±.001	.928±.002	.905±.007	.936±.004
<b>dhSegment + Page</b>	.978±.001	.929±.002	.909±.006	.938±.004
System-4.1 (MindGarage-1)	.9864	.9357	.8963	.9395
System-5 (NLPR)	.9835	.9365	.9271	.9490

Train : 150 (50/50/50)

Test : 30 (10/10/10)

[1] F. Simistira, et al., "ICDAR2017 competition on layout analysis for challenging medieval manuscripts," 2017



# Two examples on homogeneous collections

# Example : Table columns

## Homogeneous collection

Numero della Mappa	POSSessori	Denominazione del Pato di terra	QUALITA'	Classe	Superficie Pubblica Comune	Classe	Numero della Mappa
699	Marche Tommaso Giuseppe	abaz, abaz, abaz	Class. 2. off.				
700	Tommaso Antonio Tommaso	abaz, abaz, abaz	Class. 2. off. con una ditta				
701	Roberto Poligo	abaz, abaz	Class. 2. off.				
702	Luigi Carlo del G. Pietro	abaz, abaz, abaz	Class. 2. off. con una ditta				
703	Riccardo Carlo del G. Tommaso	abaz, abaz	Class. 2. off.				
704	Marche Luigi	abaz	Class. 2. off.				
705	Giuseppe Giovanni Andrea	abaz	idem				
706	Luigi Giovanni Andrea	abaz	idem				
707	Marche Luigi	abaz	idem				
708	Riccardo Andrea G. Tommaso	abaz, abaz	idem				
709	Tommaso Antonio Tommaso	abaz, abaz	idem				
710	Luigi Luigi, e Massimo Pietro	abaz, abaz	Class. con due ditta 2. off.				
711	Marche Luigi	abaz	Class. 2. off.				
712	Tommaso Antonio Tommaso	abaz, abaz	idem				

Numero della Mappa	POSSessori	Denominazione del Pato di terra	QUALITA'	Classe	Superficie Pubblica Comune	Classe	Numero della Mappa
791	Tommaso Antonio Tommaso	abaz, abaz, abaz	Class. 2. off.				
792	Luigi Luigi	abaz, abaz, abaz	Class. 2. off.				
793	Tommaso Antonio Tommaso	abaz, abaz, abaz	Class. 2. off.				
794	Roberto Poligo	abaz, abaz	Class. 2. off.				
795	Luigi Luigi, e Massimo Pietro	abaz, abaz	Class. 2. off.				
796	Luigi Luigi	abaz, abaz	Class. 2. off.				
797	Luigi Luigi	abaz, abaz	Class. 2. off.				
798	Luigi Luigi	abaz, abaz	Class. 2. off.				
799	Luigi Luigi	abaz, abaz	Class. 2. off.				
800	Luigi Luigi	abaz, abaz	Class. 2. off.				
801	Luigi Luigi	abaz, abaz	Class. 2. off.				
802	Luigi Luigi	abaz, abaz	Class. 2. off.				
803	Luigi Luigi	abaz, abaz	Class. 2. off.				
804	Luigi Luigi	abaz, abaz	Class. 2. off.				
805	Luigi Luigi	abaz, abaz	Class. 2. off.				
806	Luigi Luigi	abaz, abaz	Class. 2. off.				
807	Luigi Luigi	abaz, abaz	Class. 2. off.				
808	Luigi Luigi	abaz, abaz	Class. 2. off.				
809	Luigi Luigi	abaz, abaz	Class. 2. off.				
810	Luigi Luigi	abaz, abaz	Class. 2. off.				

Trained on 17 images  
Processed ~1400 pages



# Example : Cadaster maps



Trained on 1/3 of a map sheet (~800 parcels)

Processed 26 map sheets (> 40'000 parcels)



# Conclusion

Using the **same network** on three different tasks leads to **competitive results** with state-of-the-art methods

**Pre-trained** part of the network allows a **quick training process** as well as **little amount of training data**

Presented results use **simple post-processing**, they can be largely improved using more advanced algorithms

**Further work** should investigate how incremental or parallel training over various tasks could boost **transfer learning**

# References

1. Grüning, T., Labahn, R., Diem, M., Kleber, F., & Fiel, S., “READ-BAD: A New Dataset and Evaluation Scheme for Baseline Detection in Archival Documents”, 13th IAPR International Workshop on Document Analysis Systems, 2018.
2. Simistira, F., Seuret, M., Eichenberger, N., Garz, A., Liwicki, M., & Ingold, R., “Diva-HisDB: A precisely annotated large dataset of challenging medieval manuscripts”. *Frontiers in Handwriting Recognition (ICFHR)*, 2016 15th International Conference on. IEEE, 2016.
3. Diem, M., Kleber, F., Fiel, S., Grüning, T., & Gatos, B., “cBAD: ICDAR2017 Competition on Baseline Detection.”, *Document Analysis and Recognition (ICDAR)*, 2017 14th IAPR International Conference on. Vol. 1. IEEE, 2017.
4. Simistira, F., Bouillon, M., Seuret, M., Wursch, M., Alberti, M., Ingold, R., & Liwicki, M., “ICDAR2017 competition on layout analysis for challenging medieval manuscripts”. Simistira, Fotini, et al. "Icdar2017 competition on layout analysis for challenging medieval manuscripts." 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). IEEE, 2017.
5. Tensmeyer, C., Davis, B., Wigington, C., Lee, I., & Barrett, B. (2017, November). “Pagenet: Page boundary extraction in historical handwritten documents.” *Proceedings of the 4th International Workshop on Historical Document Imaging and Processing*. ACM, 2017.
6. Ronneberger, O., Fischer, P., & Brox, T., “U-net: Convolutional networks for biomedical image segmentation.” *International Conference on Medical image computing and computer-assisted intervention*. Springer, Cham, 2015.
7. He, K., Zhang, X., Ren, S., & Sun, J., “Deep residual learning for image recognition”. *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

# Resources and contact



[github.com/dhlab-epfl/dhSegment](https://github.com/dhlab-epfl/dhSegment)

**Digital Humanities Laboratory**  
dhlab.epfl.ch

**Sofia Ares Oliveira**  
sofia.oliveiraares@epfl.ch

**Benoit Seguin**  
benoit.seguin@epfl.ch

**Frederic Kaplan**  
frederic.kaplan@epfl.ch

