

DOCUMENT LAYOUT ANALYSIS

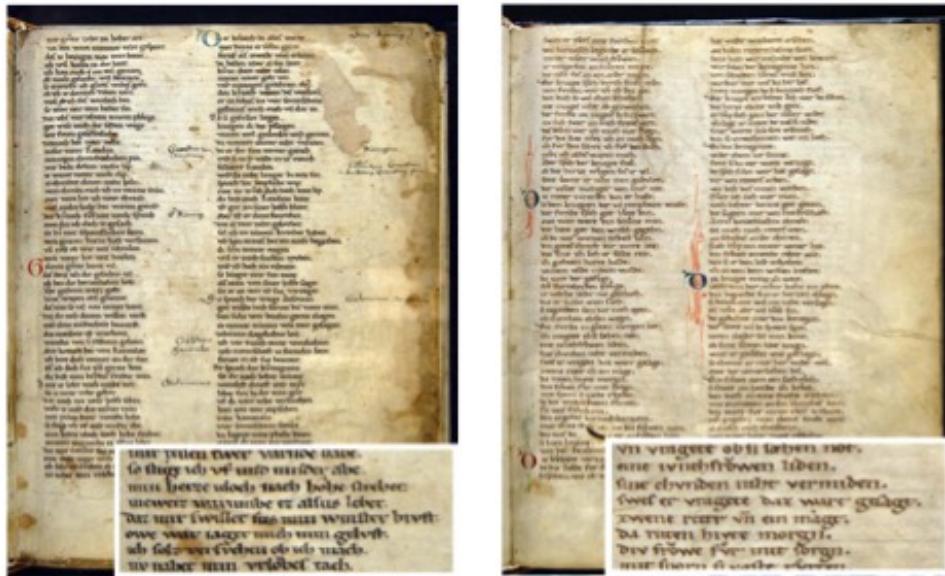
Olarik Surinta

Mahasarakham University

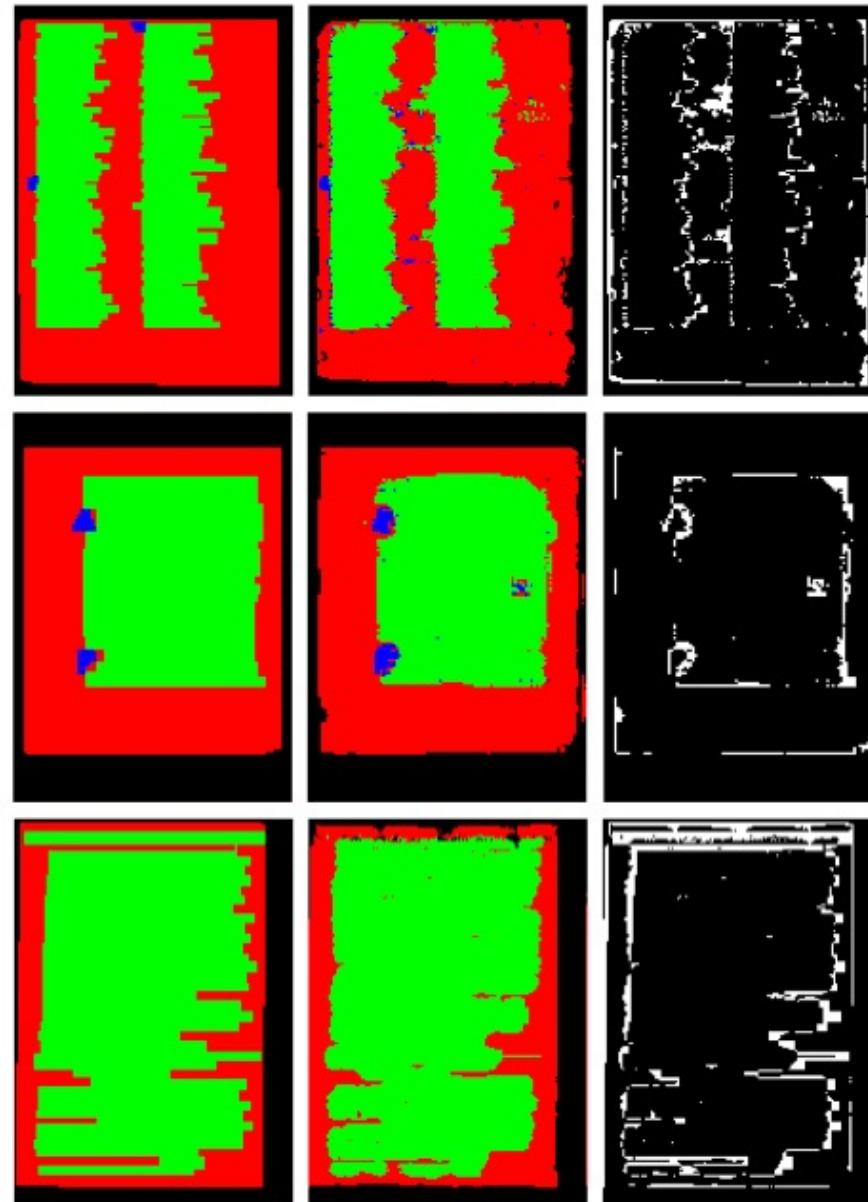
Document Layout Analysis

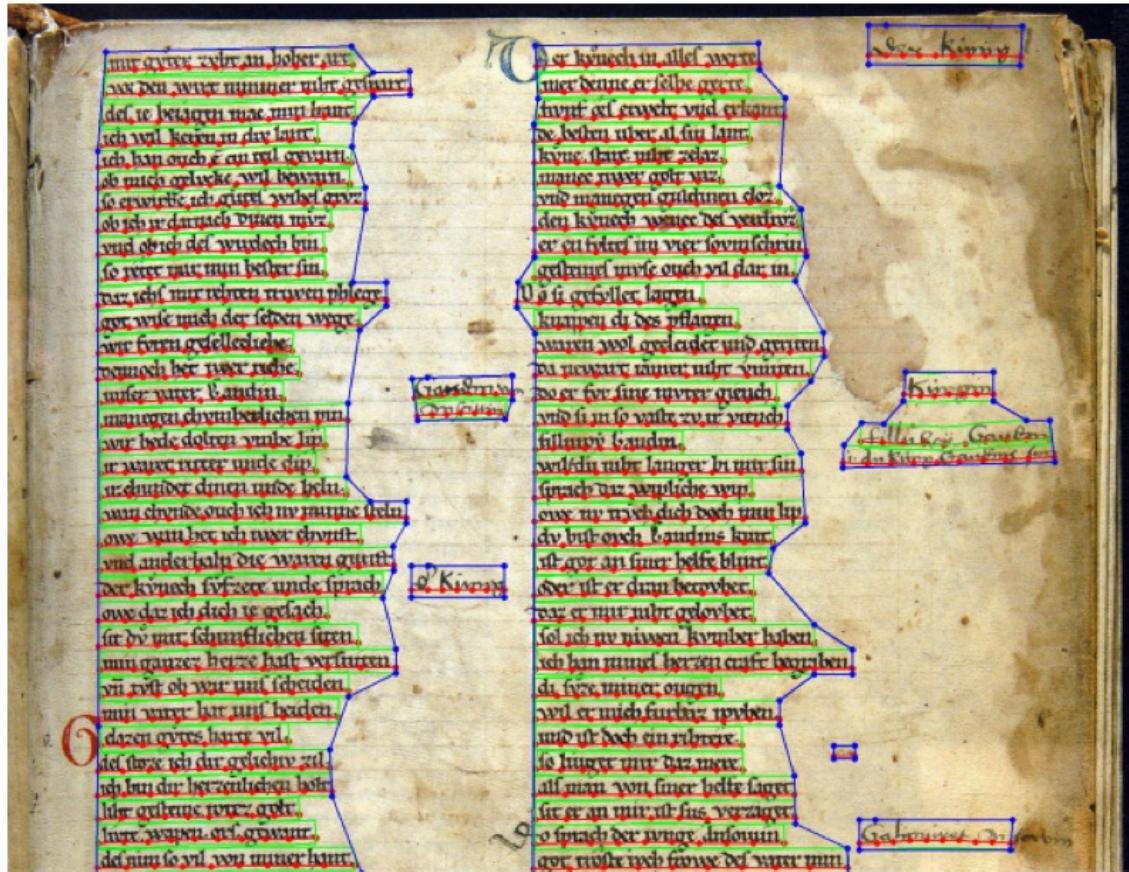
Document layout analysis (DLA) is performed to determine physical structure of a document, that is, to determine document components.

Historical Document Layout Analysis



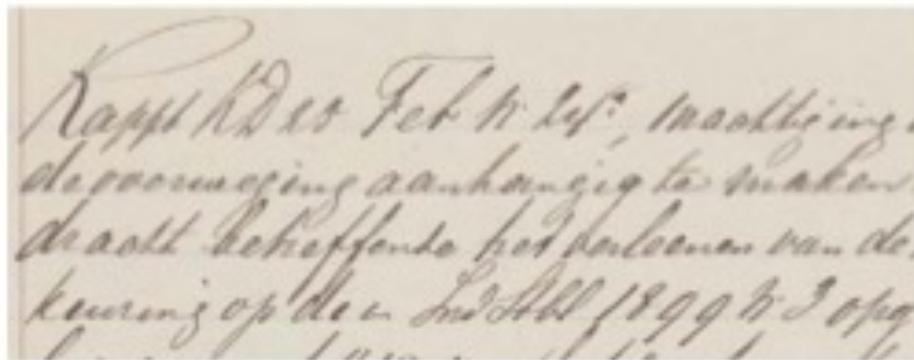
Wei et al. (2014) Hybrid feature selection for historical document layout analysis, DOI: [10.1109/ICFHR.2014.22](https://doi.org/10.1109/ICFHR.2014.22)



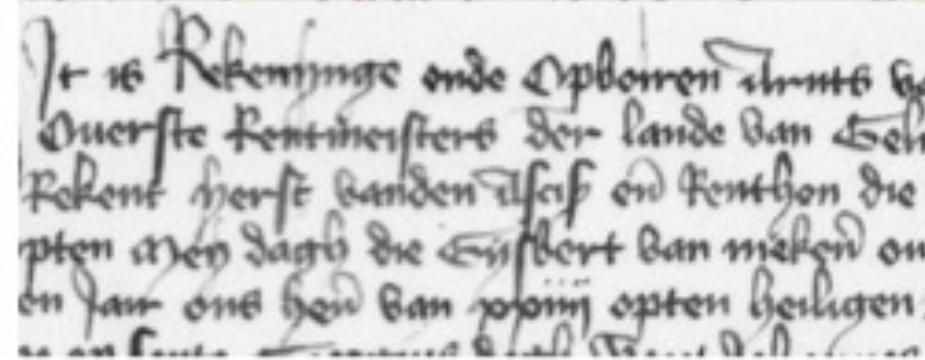


essa suscep*tus* ē ab eo cum omni honore. & gaudio magno.
Mansit itaq. apud illum aliquor diebus. sacras scripturas
aperiens. & insinuans ei veritatem. Cumq. & ipsum ro-
garo. ut adagilolfum regem langobardorum. eius iussu
per alamannum ducetur. malefice ferens rex tantoru-
m discessum uirorum. pollicitus ē ei*s* infra terminos sui regnū
se reperiturum loca uenusta. quae famulis di. & adincolen-
dum essent commoda. & admisit uendas uerba ueritatis cui
cum positas nationes oportuna; Vir di. semen uerbi incor-
dibus gentium plantare desiderant. aliquantis per moratu-
rum se promisit. si regia auctoritas factis dicta firmara;
ACEP*T*A I^UITUR A*S*CE*L*ICENTIA UIGIENI
locum ubicumq. uoluisser. dum loca plurima per-
luerassent. uenerunt infra partes alamanniae. Ad
fluum qui lindimacus vocatur; Luxta quem adsuperi-
ora tendentes. puerunt ad lacum turcianum; Cumq.
per litus ambalantes uenissent ad caput lacu*s* ipsius. In locum
qui tuconia dicitur. placuit illis loci qualitas adinhabi-
randum; Porro homines ibidem commanentes. crudeles
erant. & impu. simulacra colentes. idola sacrificiis ueneran-
tes. obseruantes auguria. & divinationes. & multa quae con-

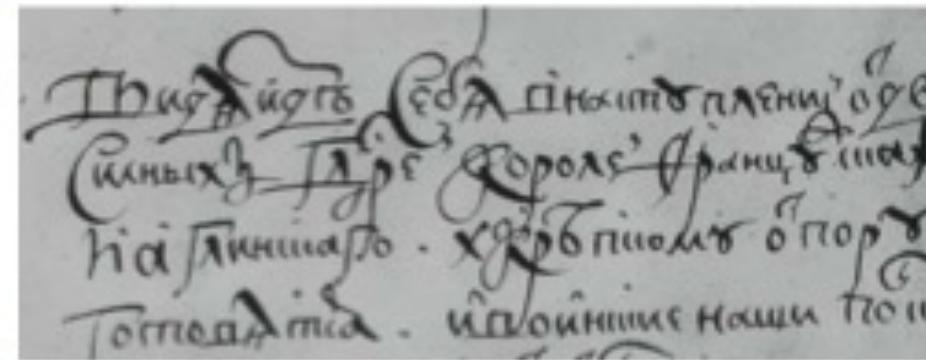
Line Segmentation



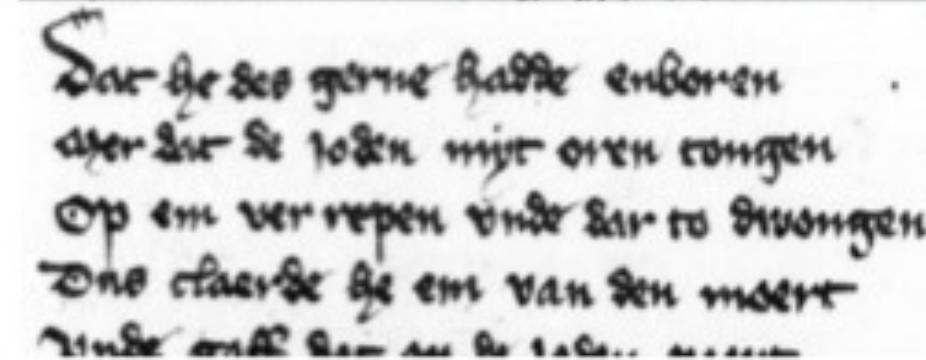
Nappt dat het wyls mochtig
de voorwaering aanhangig te maken.
daatt behoffende het volcomen van de
koning op den 1^{ste} Julij 1899, I opg.



It is Bekendinge ende Opboeren d'rents van
Ouerste Rentmeesters der lande van Ben
bekent herst banden t'self en tenthou die
opten allen dagh die Eustert van maken om
en jaer ons heu van jorom opten heiligen.

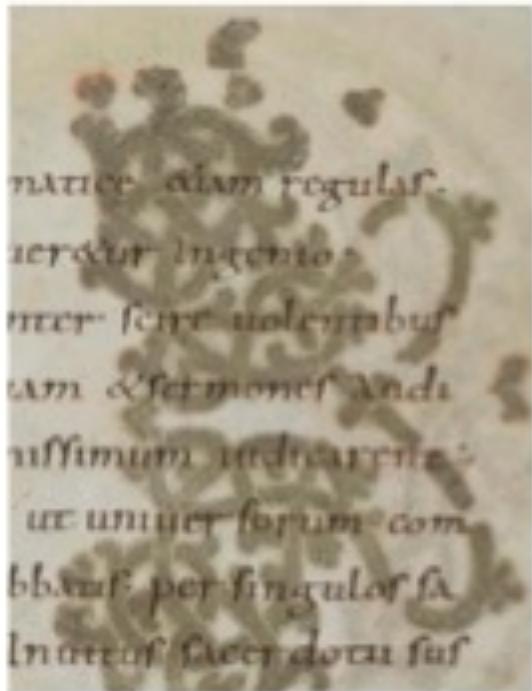


Днъ 1^{го} Сентября въ градѣ Франкфуртѣ
Синихъ Гербовъ Короля Франца ІІІ
на Ганилье. Хрестъ писанъ бѣлору
Томпсонъ. и външніе наши гон



Dat die des ghetre brachte enboven
aer dat de joden myt oren sonden
Op em verhepen vnde dat so vliedengen
Doe daerde die em van den moert
Vnde enke dan nu de eage...

Line segmentation



(A)

(B)

Binarization using (A) Otsu, (B) Sauvola algorithms

Problem of original A[★] Path Planning Algorithm

Qui nra reliquias
dim sequeremur.

(A)

- (A) Correctly separates two character lines
- (B) Cannot divide the two touching text lines.

Ende alte dont
is tformen kann.

(B)

Qui nra reliquias: ut secundum euangelicam iussionem
dm̄ sequeremur: non debemus alienas amplecti diuitias.

(a)

It is Rekeninge ende Explorren droms gen aenhop
Overste Rennemaster der lande van Sebre. Ende

(b)

Het volkinge regtder wordt ten aansien van
de Godicheit. Gerechtigheyt en Thwaertichyt by den dag

(c)

Thw. Augs. Cestia p. k. i. t. pl. 8
Синыхъ дѣпъ Родоръ дѣни

(d)

на границе . хребты огорожены
Городами . и покинуты наши посты
Сибирские посты ограблены
Птичники . а покинута Кара-Ли
Надежда у птицей склон
и . на Амуре Наморь
Сибирь Сибирь уничтожена .
птичники горело у нас (сюда) сидевшего

Socrates was a classical
credited as one of the first
philosophy, he is an enigma
only through the classical
students, Plato's dialogues

(a) The original binary image.



(b) SDT of the image in Fig. 1(a).

Fig. 1. An example of a binary image and its SDT. In SDT, the darker one point is, the lower value its SDT has.

SDT = Singed Distance Transform

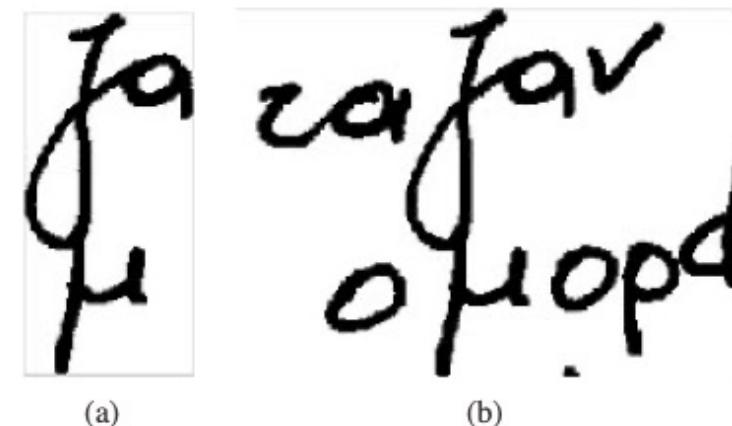


Fig. 2. A large components and its neighbouring strokes lying in the same text lines. (a) is a large component detected. In (b) are the horizontal neighbouring strokes.

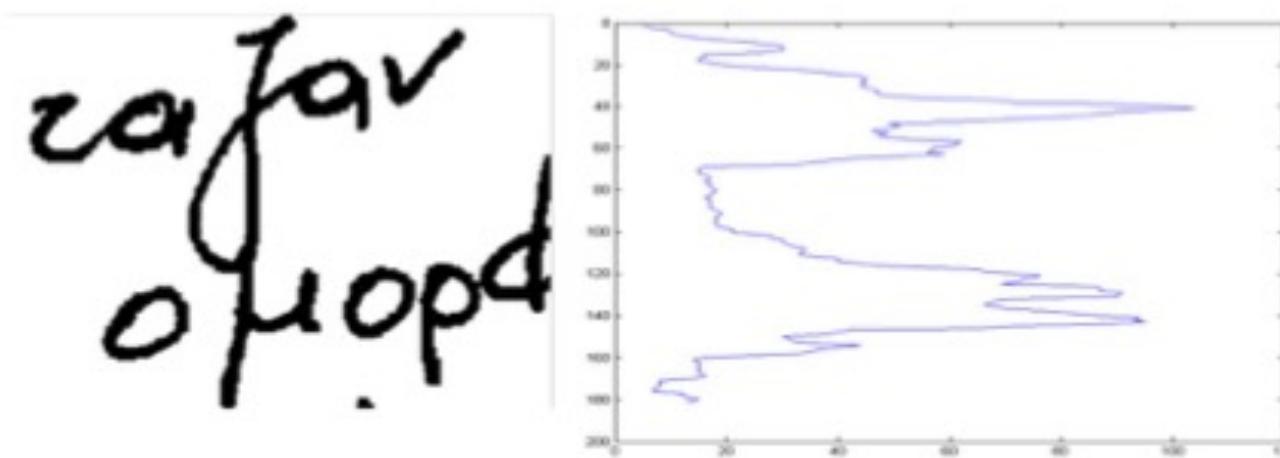


Fig. 3. The horizontal histogram projection of the image in Fig. 2(b)

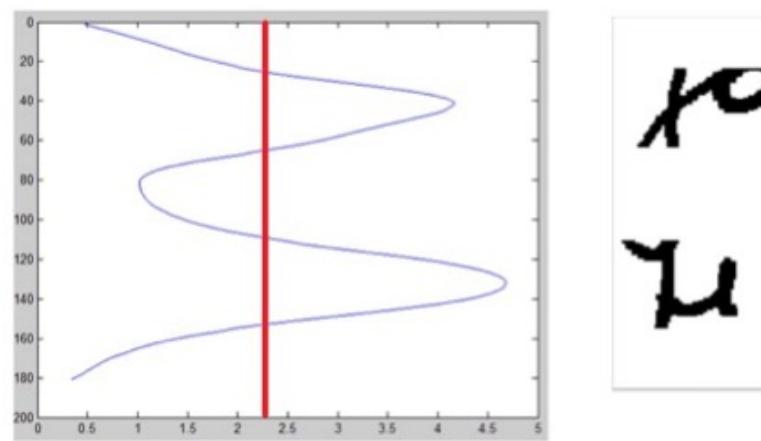


Fig. 4. Choose a threshold on the smoothed histogram, indicating by the red line, and remove the rows with lower intensities.

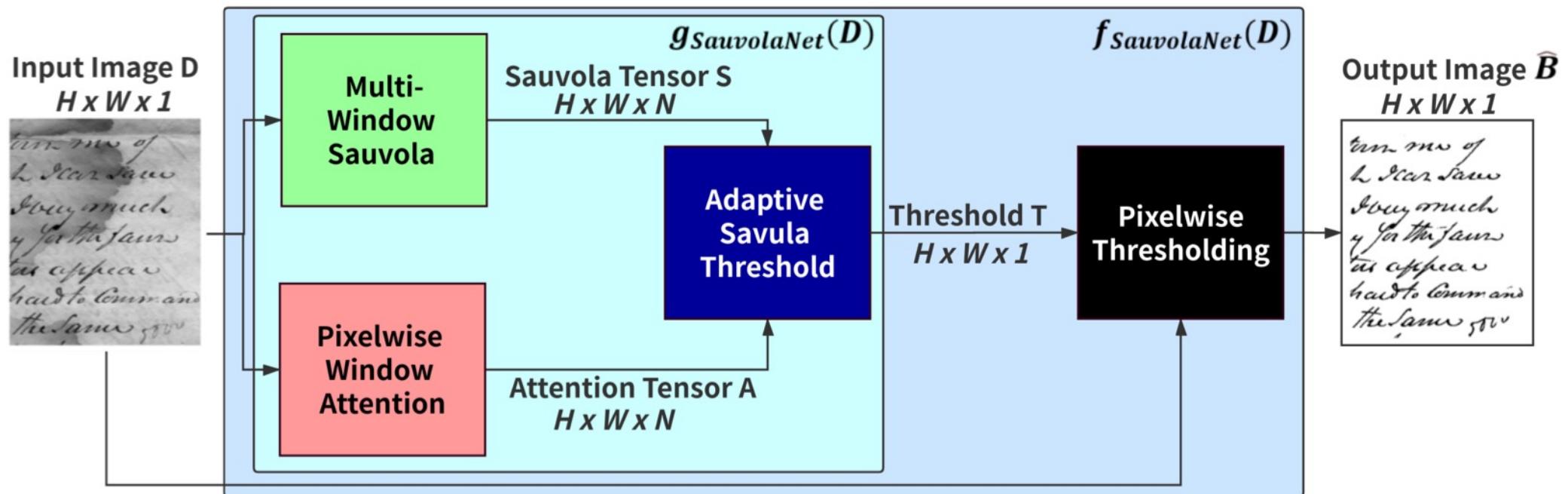
~~Socrates was a classical s₁~~
~~credited as one of the f₂~~
~~philosophy, he is an enigma s₃~~
~~only through the classical s₄~~
~~students, Plato's dialogues s₅~~

Fig. 7. The final seams detected by our proposed method. There are total five seams, indicating the central axis positions of five text lines.

ignore
sell - km

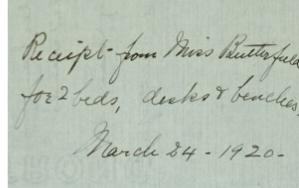
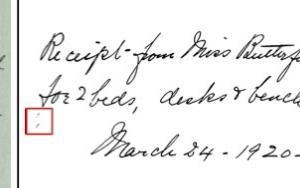
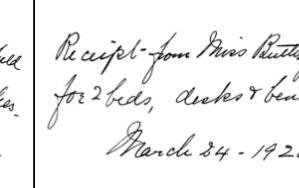
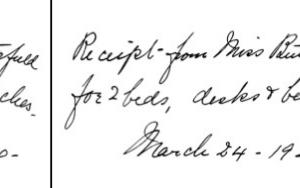
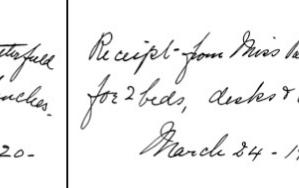
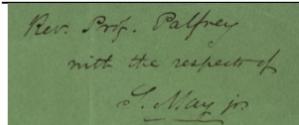
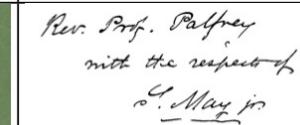
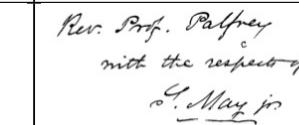
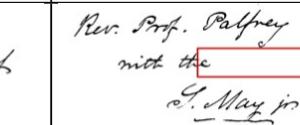
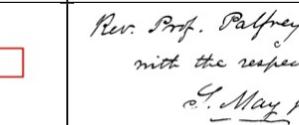
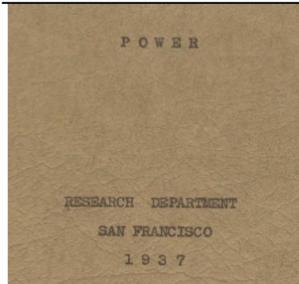
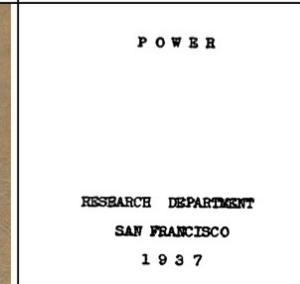
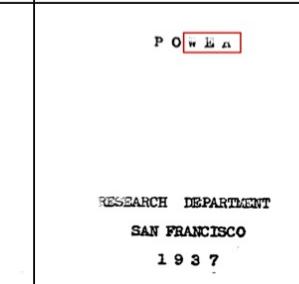
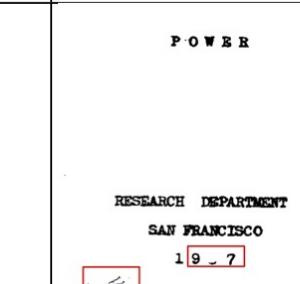
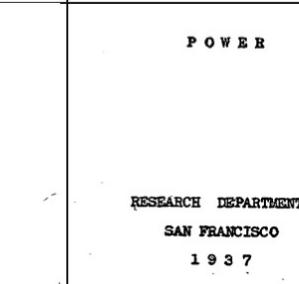
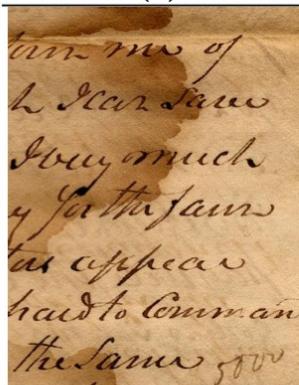
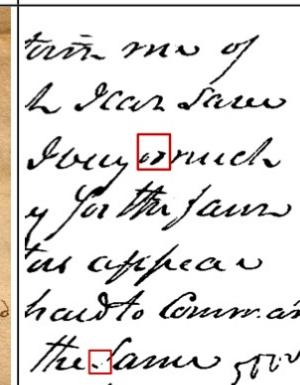
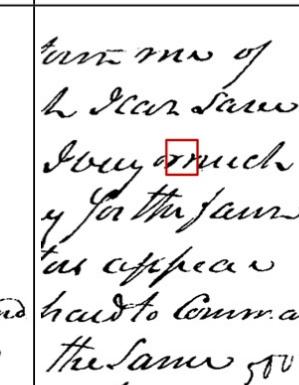
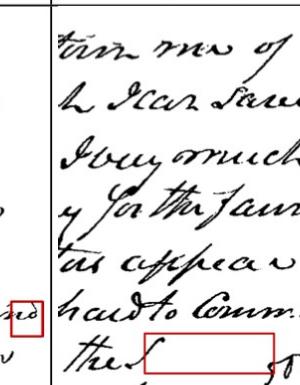
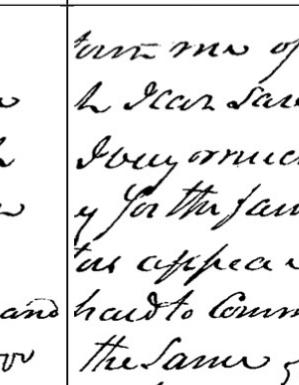
Η Ινδία, η Επικράτεια Δημοκρατία των Ινδίων είναι Χώρα στη
Βόρεια Ασία. Έχει μεγάλη προστίχεια χώρας Παγκόσμιων
ετών πληθυσμού, με 1.166.079.217 κατοίκους, και η επιφάνεια
πληθυσμού σε έκταση με 3.287.590 km². Εγκέριται
διακρίτης στη Μεράλα ή όπως γνωστή θεωρείται από τον
ονόμα αριθμού της Ινδίας, η οποία διατηρείται από την παραδοσιακή.
Συνορεύει συρρότιμα με τη Λαϊκή Σοβιετική Ένωση στη Βεγκαζίτη,
Βόρεια με την Κίνα και τη ινδο-ινδιανή Λιανική Νεταζή,
Βορειοδυτικά με τη Ρωσία, Ενώ δυτικά βρέχεται από
την Αραβική Θάλασσα και νότια, νοτιοδυτικά από την
Ινδική θάλασσα και την ισλαμική Βεγγαζίτη Ηνωμένη,
νατά στη Βιετναμική της Ιεροπόλις υπό την ονομασίαν
Εργάτη πολιτών την πατριωτική πολιτείαν ή και Υπερβολική/
Κατά τη διάρκεια των 1800 άνετα οριστικά υποτάχθηκε
την Ινδιανή Βασίλειο, του το 1947 απευθύνεται
επεξαργήσια με βάση την γεωγραφία αριθμού παραδοσιακής
που ξαραναπίστηκε από την Βιετναμική πλευράς. Σήμερα
η Ινδία έχει από την πλευρά παραγγελμάτων
οικονομίες, πολιτική και στρατιωτική παρέμβει από την
κοινωνία οικολογία.

SauvolaNet: Learning Adaptive Sauvola Network for Degraded Document Binarization

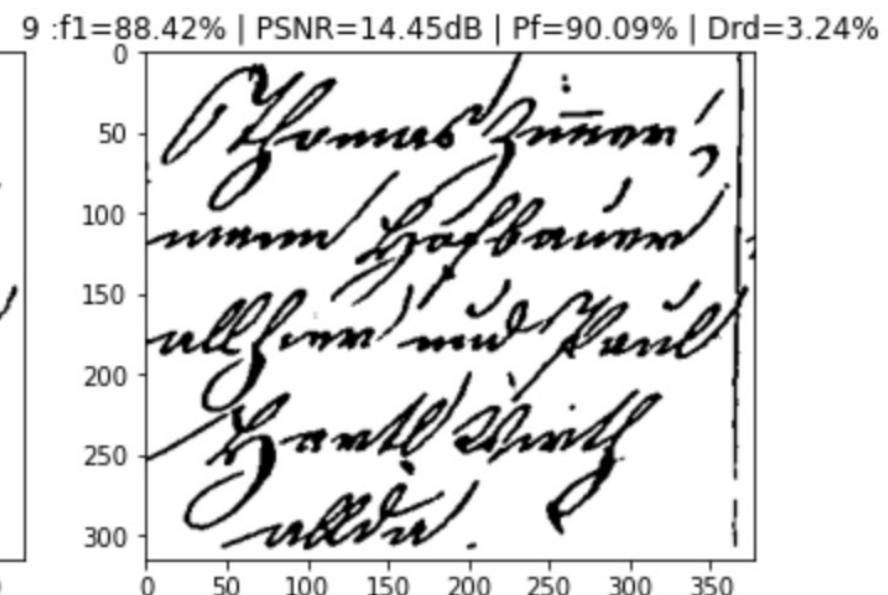
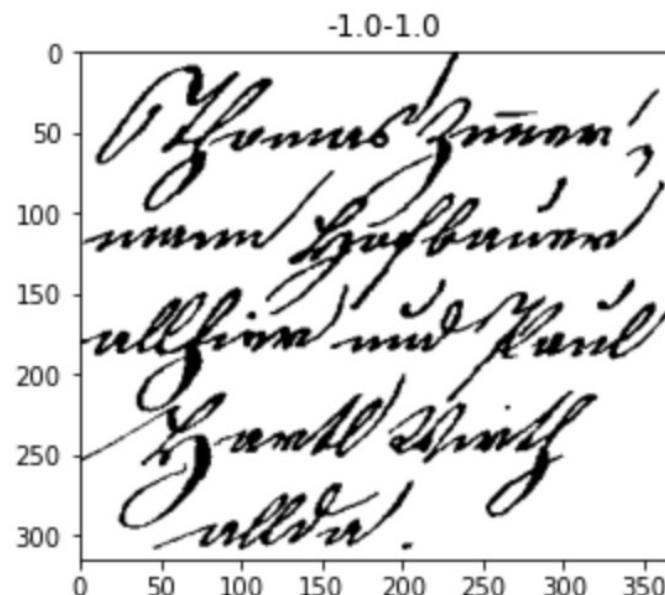
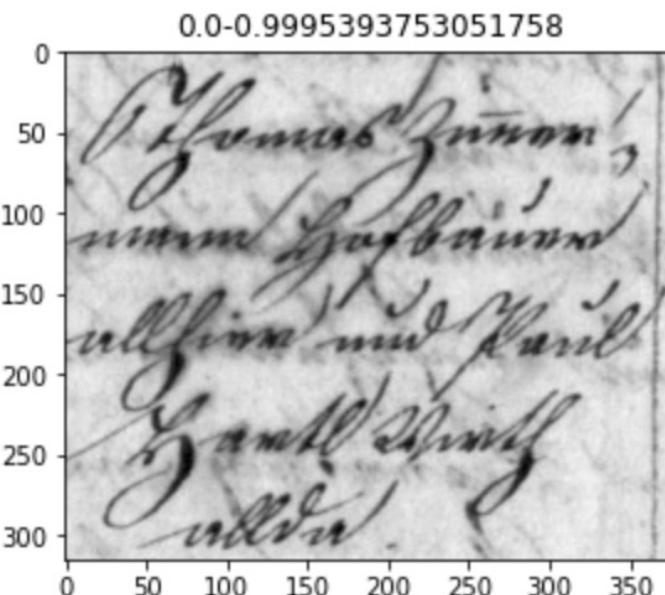


SauvolaNet: <https://arxiv.org/abs/2105.05521>

<https://github.com/Leedeng/SauvolaNet>

Original	cGANs [34]	DeepOtsu [9]	MRAtt [19]	SauvolaNet
	 March 34 - 1920-	 March 34 - 1920-	 March 34 - 1920-	 March 34 - 1920-
(a)	95.89	96.28	95.91	97.81
	 G. May Jr.	 G. May Jr.	 G. May Jr.	 G. May Jr.
(b)	93.79	95.86	86.01	96.64
				
(c)	91.89	88.48	90.25	93.85
	 the same go	 the same go	 the same go	 the same go
(d)	91.98	91.73	91.13	94.37

SauvolaNet:



Form Analysis



Page 2

FACTURE

Date : 21/12/2012

N° : 00213969

34, avenue du soleil levant
03 96 26 35 74
03 96 26 35 35
hallstarindus@bmail.com
www.hall-star.com

GAUTHIER Laurent
98, rue Jacques Callot
03 87 34 43 44

N° client	N° dossier
0011263	0011325

Poirriez et al. (2014) Visual perception of unitary elements for layout analysis of unconstrained documents in heterogeneous databases, DOI: [10.1109/ICFHR.2014.14](https://doi.org/10.1109/ICFHR.2014.14)

La main d'œuvre de nos travaux et entretiens est valorisée FORFAITAIREMENT suivant barèmes des temps forfaitsaires.

Notre organisation et gestion sont certifiées ISO 9001



Table analysis

Quantité	Article	Unités	Description	Réduc %	H.T.	Prix unitaire	Total
4	R19	20	Stylos	/		4€	80€
12	T73	5	Celle VHO	/		8€	40€
24	S66	10	Gomme, bâton	/		2,90€	29,00€
100	U72	1	Compas	/		6€	600
101	C88	1	Équerres	/		6,5€	656,5
						Sous-total	1047,10
						I.P.S.	
						I.V.Q.	
						Livraison	
						Montant à verser	1047,10



Poirriez et al. (2014) Visual perception of unitary elements for layout analysis of unconstrained documents in heterogeneous databases, DOI: [10.1109/ICFHR.2014.14](https://doi.org/10.1109/ICFHR.2014.14)

Cosmic String Detection with Tree-Based Machine Learning

A. Vafaei Sadj^{1,2,3}, M. Forhangi¹, S. M. S. Moosavi^{1,2,4}, B. Bassiri^{1,2,4,5}, M. Komei²
¹ Department of Physics, Shahid Beheshti University, Gorgan, Tehran, Iran
² Department of Physics, Shahrood University of Technology, Shahrood, Iran
³ Institute for Research in Theoretical Physics & Mathematics, 4 Bahman Blvd, Ashtekar 19741, South Africa
⁴ ATLAS Experiment at CERN, 2025 Route de Meyrin, 1211 Geneva 15, Switzerland
⁵ School of Physics, Institute for Research in Theoretical Physics, P.O. Box 19395-5747, Tehran, Iran
⁶ South African Astronomical Observatory, Observatory, Cape Town, 7400, South Africa
⁷ Department of Physics, The University of Port Elizabeth, Port Elizabeth, South Africa
⁸ Department of Maths and Applied Maths, University of Cape Town, Cape Town, South Africa

21 November 2019

ABSTRACT

We explore the use of random forest and gradient boosting tree-based machine learning algorithms for the detection of cosmic strings in maps of the cosmic microwave background (CMB) through their unique features that are different from the noise and other artifacts. The reference map is composed into four main sectors before being passed to the learning units. The feature vectors contain various statistical properties of the input maps, such as the mean, standard deviation, and correlation coefficient. Our proposed classifiers, after training, give results improved over or similar to the claimed detectability levels of the existing methods for string detection. They can also detect strings with a length of up to 100 pixels and a density of up to 10^{-3} pixels per observation. The minimum detectable density increases to $\sim 1.2 \times 10^{-3}$ for the tree models, CMB WMAP 9 (II) strategy, with a significant improvement over the previous results.

Key words: Cosmic string, Machine learning, Tree based model, CMB, WMAP.

1. INTRODUCTION

The inflationary paradigm is the most widely accepted scenario for solving the anomalies in the Universe, so far presented. This scenario predicts that the early Universe, both theoretical and observational, can form before inflation, with different densities. Among other predictions, it predicts the formation of topological defects formed as a result of phase transitions. In particular, cosmic strings (CS) are considered to be the predicted defect of the inflationary paradigm (Vilenkin 1986; Vilenkin 1989; Vachaspati & Vilenkin 1995; Vilenkin 1995; Michael 1997; Michaelis & Riddle 1998; Vilenkin & Shellard 1994; Vilenkin & Shellard 1994; Vilenkin & Shellard 1994; Davis et al. 2000; Copeland et al. 2004; Andreassen et al. 2005; Davis et al. 2005; Davis et al. 2006; Davis & Vilenkin 2006; Kilbey & Davis 2006). The detection of CS would be a major breakthrough in cosmology. The first detections of CS were made by the COBE mission (Mather et al. 1990; Lubin & Jaffe 1990; Vittorio 1990; Vilenkin & Michael 1990; Penitenti & Tye 2000). There has been a lot of work on the detection of CS in the literature (e.g., Vilenkin & Shellard 1994; Vilenkin 1995; Vilenkin & Shellard 1994; Davis et al. 2000; Copeland et al. 2004; Andreassen et al. 2005; Davis et al. 2005; Davis et al. 2006; Davis & Vilenkin 2006; Kilbey & Davis 2006).

1.1. Related Work

After low-level but interesting powerful statistical tools for cosmic string detection and pattern recognition, several methods have been proposed for the detection of CS and a rigorous theoretical analysis of the string sources, respectively. The string sources is intimately related to the origin of the phase transition spots.

$$\frac{d\sigma}{d\Omega} = \frac{1}{2} \left(\frac{\partial^2 \mathcal{L}}{\partial \theta^2} \right)_{\text{min}}$$

where \mathcal{L} is the Lagrangian function and θ is the angle between the direction of the string and the direction of the momentum. The parameters are used to control units with $1 \times 1 \times 1$. A 1D network would have serious impacts on the cost of computation and the number of parameters. Therefore, we propose a 2D network to detect strings in the input maps. We use the same architecture as the 1D network, but the input size is 100×100 . The output of the network is a binary map with a size of 100×100 . The output of the network is a binary map with a size of 100×100 .

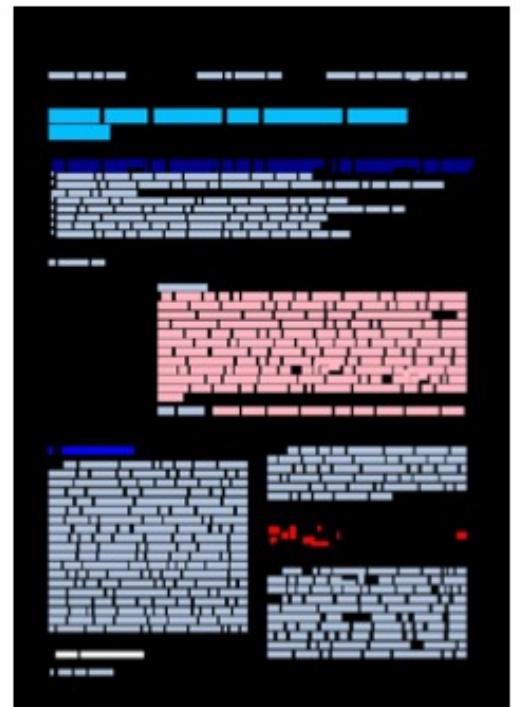


Figure 1: Example annotations of the DocBank. The colors of semantic structure labels are: **Abstract** (pink), **Author** (blue), **Caption** (green), **Equation** (red), **Figure** (yellow), **Footer** (purple), **List** (dark blue), **Paragraph** (light green), **Reference** (orange), **Section** (dark green), **Table** (cyan), **Title** (light blue).

Li et al. (2020) DocBank: A Benchmark Dataset for Document Layout Analysis, <https://arxiv.org/pdf/2006.01038v3.pdf>

Cosmic String Detection with Tree-Based Machine Learning

5

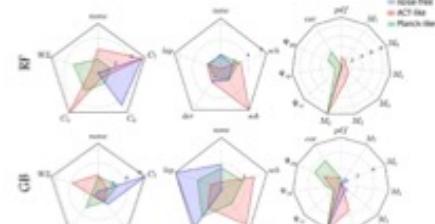


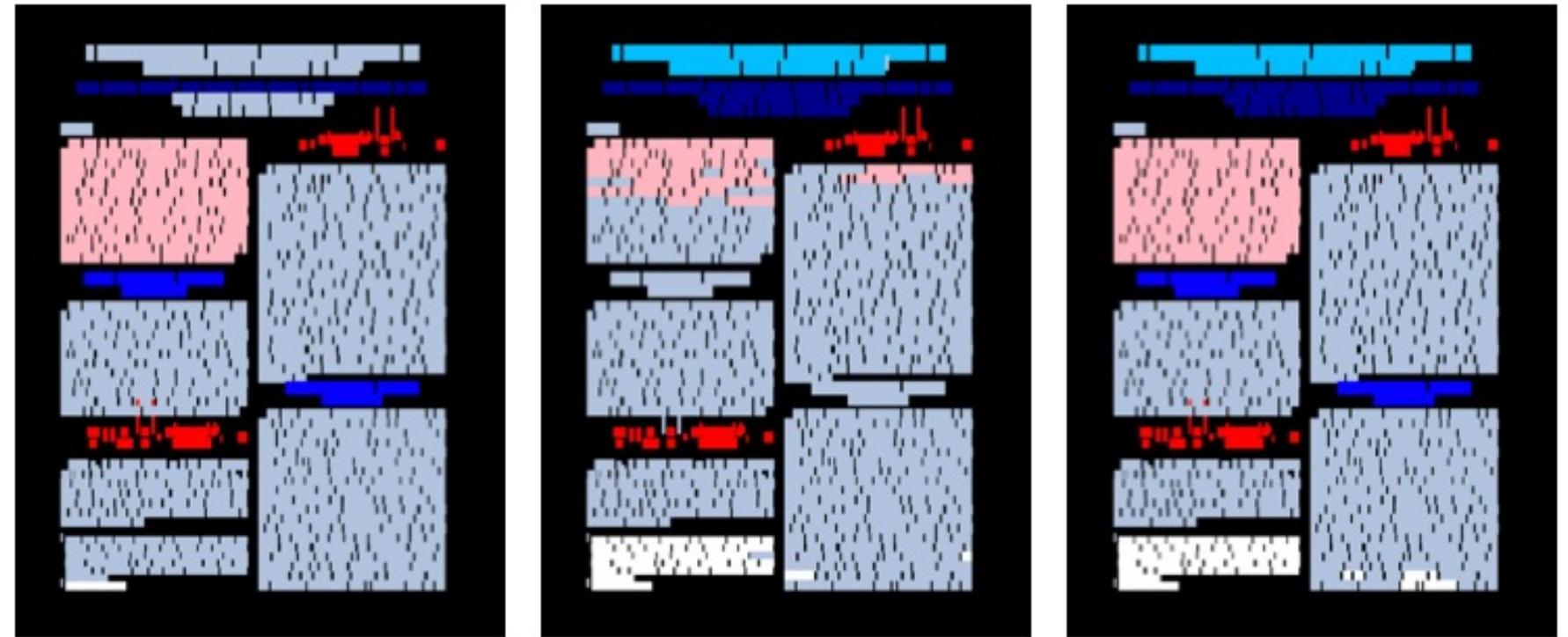
Figure 2: Feature importance report: the average number of times each feature appeared among the top ten features, normalized by the performance for the RF (blue) and GB (orange) classifiers.

Table 6: The minimum detectable Ω_{CS} in Ω_{CDM} for the three machine learning algorithms.

experiment	$\Omega_{\text{CS}}(\text{RF})$	$\Omega_{\text{CS}}(\text{GB})$
noise-free	3.1×10^{-3}	3.1×10^{-3}
CMB WMAP 9 (I)	3.1×10^{-3}	3.1×10^{-3}
CMB WMAP 9 (II)	3.1×10^{-3}	3.1×10^{-3}
MLP	3.1×10^{-3}	3.1×10^{-3}
Phenom D	3.1×10^{-3}	3.1×10^{-3}

Table 6 shows the minimum detectable Ω_{CS} in Ω_{CDM} for the three machine learning algorithms. The noise-free case is 1.8 times larger than the case of the CMB WMAP 9 (II) strategy. The noise-free case is 1.8 times larger than the case of the CMB WMAP 9 (II) strategy. The noise-free case is 1.8 times larger than the case of the CMB WMAP 9 (II) strategy. The noise-free case is 1.8 times larger than the case of the CMB WMAP 9 (II) strategy.

Note that Table 6 only reports the minimum detectable Ω_{CS} and not their associated errors. That is because the errors of the three machine learning algorithms are not available. We performed feature analysis on the feature vectors to find the significance of the role of each feature in the classification process. We used the t -test to reduce the computational cost of feature analysis by decreasing the number of the feature space and keeping the most significant features. We found that the most significant features are the μ_0 , μ_1 , μ_2 , μ_3 , μ_4 , μ_5 , μ_6 , μ_7 , μ_8 , μ_9 , μ_{10} , μ_{11} , μ_{12} , μ_{13} , μ_{14} , μ_{15} , μ_{16} , μ_{17} , μ_{18} , μ_{19} , μ_{20} , μ_{21} , μ_{22} , μ_{23} , μ_{24} , μ_{25} , μ_{26} , μ_{27} , μ_{28} , μ_{29} , μ_{30} , μ_{31} , μ_{32} , μ_{33} , μ_{34} , μ_{35} , μ_{36} , μ_{37} , μ_{38} , μ_{39} , μ_{40} , μ_{41} , μ_{42} , μ_{43} , μ_{44} , μ_{45} , μ_{46} , μ_{47} , μ_{48} , μ_{49} , μ_{50} , μ_{51} , μ_{52} , μ_{53} , μ_{54} , μ_{55} , μ_{56} , μ_{57} , μ_{58} , μ_{59} , μ_{60} , μ_{61} , μ_{62} , μ_{63} , μ_{64} , μ_{65} , μ_{66} , μ_{67} , μ_{68} , μ_{69} , μ_{70} , μ_{71} , μ_{72} , μ_{73} , μ_{74} , μ_{75} , μ_{76} , μ_{77} , μ_{78} , μ_{79} , μ_{80} , μ_{81} , μ_{82} , μ_{83} , μ_{84} , μ_{85} , μ_{86} , μ_{87} , μ_{88} , μ_{89} , μ_{90} , μ_{91} , μ_{92} , μ_{93} , μ_{94} , μ_{95} , μ_{96} , μ_{97} , μ_{98} , μ_{99} , μ_{100} , μ_{101} , μ_{102} , μ_{103} , μ_{104} , μ_{105} , μ_{106} , μ_{107} , μ_{108} , μ_{109} , μ_{110} , μ_{111} , μ_{112} , μ_{113} , μ_{114} , μ_{115} , μ_{116} , μ_{117} , μ_{118} , μ_{119} , μ_{120} , μ_{121} , μ_{122} , μ_{123} , μ_{124} , μ_{125} , μ_{126} , μ_{127} , μ_{128} , μ_{129} , μ_{130} , μ_{131} , μ_{132} , μ_{133} , μ_{134} , μ_{135} , μ_{136} , μ_{137} , μ_{138} , μ_{139} , μ_{140} , μ_{141} , μ_{142} , μ_{143} , μ_{144} , μ_{145} , μ_{146} , μ_{147} , μ_{148} , μ_{149} , μ_{150} , μ_{151} , μ_{152} , μ_{153} , μ_{154} , μ_{155} , μ_{156} , μ_{157} , μ_{158} , μ_{159} , μ_{160} , μ_{161} , μ_{162} , μ_{163} , μ_{164} , μ_{165} , μ_{166} , μ_{167} , μ_{168} , μ_{169} , μ_{170} , μ_{171} , μ_{172} , μ_{173} , μ_{174} , μ_{175} , μ_{176} , μ_{177} , μ_{178} , μ_{179} , μ_{180} , μ_{181} , μ_{182} , μ_{183} , μ_{184} , μ_{185} , μ_{186} , μ_{187} , μ_{188} , μ_{189} , μ_{190} , μ_{191} , μ_{192} , μ_{193} , μ_{194} , μ_{195} , μ_{196} , μ_{197} , μ_{198} , μ_{199} , μ_{200} , μ_{201} , μ_{202} , μ_{203} , μ_{204} , μ_{205} , μ_{206} , μ_{207} , μ_{208} , μ_{209} , μ_{210} , μ_{211} , μ_{212} , μ_{213} , μ_{214} , μ_{215} , μ_{216} , μ_{217} , μ_{218} , μ_{219} , μ_{220} , μ_{221} , μ_{222} , μ_{223} , μ_{224} , μ_{225} , μ_{226} , μ_{227} , μ_{228} , μ_{229} , μ_{230} , μ_{231} , μ_{232} , μ_{233} , μ_{234} , μ_{235} , μ_{236} , μ_{237} , μ_{238} , μ_{239} , μ_{240} , μ_{241} , μ_{242} , μ_{243} , μ_{244} , μ_{245} , μ_{246} , μ_{247} , μ_{248} , μ_{249} , μ_{250} , μ_{251} , μ_{252} , μ_{253} , μ_{254} , μ_{255} , μ_{256} , μ_{257} , μ_{258} , μ_{259} , μ_{260} , μ_{261} , μ_{262} , μ_{263} , μ_{264} , μ_{265} , μ_{266} , μ_{267} , μ_{268} , μ_{269} , μ_{270} , μ_{271} , μ_{272} , μ_{273} , μ_{274} , μ_{275} , μ_{276} , μ_{277} , μ_{278} , μ_{279} , μ_{280} , μ_{281} , μ_{282} , μ_{283} , μ_{284} , μ_{285} , μ_{286} , μ_{287} , μ_{288} , μ_{289} , μ_{290} , μ_{291} , μ_{292} , μ_{293} , μ_{294} , μ_{295} , μ_{296} , μ_{297} , μ_{298} , μ_{299} , μ_{300} , μ_{301} , μ_{302} , μ_{303} , μ_{304} , μ_{305} , μ_{306} , μ_{307} , μ_{308} , μ_{309} , μ_{310} , μ_{311} , μ_{312} , μ_{313} , μ_{314} , μ_{315} , μ_{316} , μ_{317} , μ_{318} , μ_{319} , μ_{320} , μ_{321} , μ_{322} , μ_{323} , μ_{324} , μ_{325} , μ_{326} , μ_{327} , μ_{328} , μ_{329} , μ_{330} , μ_{331} , μ_{332} , μ_{333} , μ_{334} , μ_{335} , μ_{336} , μ_{337} , μ_{338} , μ_{339} , μ_{340} , μ_{341} , μ_{342} , μ_{343} , μ_{344} , μ_{345} , μ_{346} , μ_{347} , μ_{348} , μ_{349} , μ_{350} , μ_{351} , μ_{352} , μ_{353} , μ_{354} , μ_{355} , μ_{356} , μ_{357} , μ_{358} , μ_{359} , μ_{360} , μ_{361} , μ_{362} , μ_{363} , μ_{364} , μ_{365} , μ_{366} , μ_{367} , μ_{368} , μ_{369} , μ_{370} , μ_{371} , μ_{372} , μ_{373} , μ_{374} , μ_{375} , μ_{376} , μ_{377} , μ_{378} , μ_{379} , μ_{380} , μ_{381} , μ_{382} , μ_{383} , μ_{384} , μ_{385} , μ_{386} , μ_{387} , μ_{388} , μ_{389} , μ_{390} , μ_{391} , μ_{392} , μ_{393} , μ_{394} , μ_{395} , μ_{396} , μ_{397} , μ_{398} , μ_{399} , μ_{400} , μ_{401} , μ_{402} , μ_{403} , μ_{404} , μ_{405} , μ_{406} , μ_{407} , μ_{408} , μ_{409} , μ_{410} , μ_{411} , μ_{412} , μ_{413} , μ_{414} , μ_{415} , μ_{416} , μ_{417} , μ_{418} , μ_{419} , μ_{420} , μ_{421} , μ_{422} , μ_{423} , μ_{424} , μ_{425} , μ_{426} , μ_{427} , μ_{428} , μ_{429} , μ_{430} , μ_{431} , μ_{432} , μ_{433} , μ_{434} , μ_{435} , μ_{436} , μ_{437} , μ_{438} , μ_{439} , μ_{440} , μ_{441} , μ_{442} , μ_{443} , μ_{444} , μ_{445} , μ_{446} , μ_{447} , μ_{448} , μ_{449} , μ_{450} , μ_{451} , μ_{452} , μ_{453} , μ_{454} , μ_{455} , μ_{456} , μ_{457} , μ_{458} , μ_{459} , μ_{460} , μ_{461} , μ_{462} , μ_{463} , μ_{464} , μ_{465} , μ_{466} , μ_{467} , μ_{468} , μ_{469} , μ_{470} , μ_{471} , μ_{472} , μ_{473} , μ_{474} , μ_{475} , μ_{476} , μ_{477} , μ_{478} , μ_{479} , μ_{480} , μ_{481} , μ_{482} , μ_{483} , μ_{484} , μ_{485} , μ_{486} , μ_{487} , μ_{488} , μ_{489} , μ_{490} , μ_{491} , μ_{492} , μ_{493} , μ_{494} , μ_{495} , μ_{496} , μ_{497} , μ_{498} , μ_{499} , μ_{500} , μ_{501} , μ_{502} , μ_{503} , μ_{504} , μ_{505} , μ_{506} , μ_{507} , μ_{508} , μ_{509} , μ_{510} , μ_{511} , μ_{512} , μ_{513} , μ_{514} , μ_{515} , μ_{516} , μ_{517} , μ_{518} , μ_{519} , μ_{520} , μ_{521} , μ_{522} , μ_{523} , μ_{524} , μ_{525} , μ_{526} , μ_{527} , μ_{528} , μ_{529} , μ_{530} , μ_{531} , μ_{532} , μ_{533} , μ_{534} , μ_{535} , μ_{536} , μ_{537} , μ_{538} , μ_{539} , μ_{540} , μ_{541} , μ_{542} , μ_{543} , μ_{544} , μ_{545} , μ_{546} , μ_{547} , μ_{548} , μ_{549} , μ_{550} , μ_{551} , μ_{552} , μ_{553} , μ_{554} , μ_{555} , μ_{556} , μ_{557} , μ_{558} , μ_{559} , μ_{560} , μ_{561} , μ_{562} , μ_{563} , μ_{564} , μ_{565} , μ_{566} , μ_{567} , μ_{568} , μ_{569} , μ_{570} , μ_{571} , μ_{572} , μ_{573} , μ_{574} , μ_{575} , μ_{576} , μ_{577} , μ_{578} , μ_{579} , μ_{580} , μ_{581} , μ_{582} , μ_{583} , μ_{584} , μ_{585} , μ_{586} , μ_{587} , μ_{588} , μ_{589} , μ_{590} , μ_{591} , μ_{592} , μ_{593} , μ_{594} , μ_{595} , μ_{596} , μ_{597} , μ_{598} , μ_{599} , μ_{600} , μ_{601} , μ_{602} , μ_{603} , μ_{604} , μ_{605} , μ_{606} , μ_{607} , μ_{608} , μ_{609} , μ_{610} , μ_{611} , μ_{612} , μ_{613} , μ_{614} , μ_{615} , μ_{616} , μ_{617} , μ_{618} , μ_{619} , μ_{620} , μ_{621} , μ_{622} , μ_{623} , μ_{624} , μ_{625} , μ_{626} , μ_{627} , μ_{628} , μ_{629} , μ_{630} , μ_{631} , μ_{632} , μ_{633} , μ_{634} , μ_{635} , μ_{636} , μ_{637} , μ_{638} , μ_{639} , μ_{640} , μ_{641} , μ_{642} , μ_{643} , μ_{644} , μ_{645} , μ_{646} , μ_{647} , μ_{648} , μ_{649} , μ_{650} , μ_{651} , μ_{652} , μ_{653} , μ_{654} , μ_{655} , μ_{656} , μ_{657} , μ_{658} , μ_{659} , μ_{660} , μ_{661} , μ_{662} , μ_{663} , μ_{664} , μ_{665} , μ_{666} , μ_{667} , μ_{668} , μ_{669} , μ_{670} , μ_{671} , μ_{672} , μ_{673} , μ_{674} , μ_{675} , μ_{676} , μ_{677} , μ_{678} , μ_{679} , μ_{680} , μ_{681} , μ_{682} , μ_{683} , μ_{684} , μ_{685} , μ_{686} , μ_{687} , μ_{688} , μ_{689} , μ_{690} , μ_{691} , μ_{692} , μ_{693} , μ_{694} , μ_{695} , μ_{696} , μ_{697} , μ_{698} , μ_{699} , μ_{700} , μ_{701} , μ_{702} , $\mu_{$



(a) Original document page

(b) Groundtruth

(c) Pre-trained BERT

(d) Pre-trained LayoutLM

Figure 4: Example output of pre-trained LayoutLM and pre-trained BERT on the test set

paragraph 1005: applicable, we argue that first, the model should be able to learn more sophisticated features than other – it can successfully capture the intrinsic memory contained in the input image by learning local and global structures in the target image, without requiring domain knowledge over user input. Further, to support the downstream medical applications, the inference process for a target image should be able to generate images that are appropriate for specific or local users, so that the user can immediately obtain the desired output without waiting for several hours. Third, a personalized padding signal usually comes in different forms – it could either be an entire style transfer or a specific transformation of a single descriptive feature (e.g. the eye in the user’s face). It is given a speech instruction – therefore, it is preferable that the model processes the capability of receiving arbitrary padding signal or multiple modulus.

A variety of techniques have been developed as well as several novel research directions. Existing approaches, such as [4, 9, 10, 11, 12, 13], mainly focus on training transformation neural networks that can respond to a small set of padding signals, such as a few specific image transformations. Using these learned transformation neural networks, each of them is only used to one specific padding signal (e.g., a simple style image), and cannot generalize to unseen types specified by users, unless constrained as many networks to share the learned parameters. In contrast, our model can handle multiple padding signals simultaneously and predict corresponding results according to specific styles. Although the most recent method by [1] can process images (modality) in real-time, it has to train a single network for each specific type of transformation (e.g., a specific style image or specific transformation). In addition, it is difficult to handle multiple padding signals or other forms of padding signals without retraining the model for every type, which usually takes several hours and processes those hours being scaled to real-world applications. Our model can handle multiple padding signals within a single network, however, because their model focuses on increasing the diversity of output images and are still unable to handle diverse and complex padding signals due to the limitation (e.g., one image at a time).

In this paper, we make the first attempt to explore the real-time multi-style image manipulation task, to our best knowledge. This task is challenging since the model should be able to exploit and transform diverse and complex patterns from one specific padding signal into transformation patterns and predict corresponding results in real-time. To this end, we propose a novel *Faster R-CNN* for image manipulation (FRCNN) that constitutes a parameter network (PNet) and an image transformation network (TNets) in end-to-end manner. When it is a generic model, it predicts the diversity of key transformations that can be applied, which then takes these generated parameters, combines with its own signal generation parameters, to generate a new image. In the sense of image style transfer, the PNNet can extract any style image from the input image and then the TNNet can transform the content image to a stylized image. We argue that FRCNN can digitize over 20,000 style images in a single second rather than training one network per style as we need previous methods did. It can also be used to process general images, as well as medical images, such as radiographs, tomographs,

paragraph 995: and even non-medical images. The proposed FRCNN can generate combination of a single image or real-time frame of soft-tissue and image crop from the input image and produce the output image in real-time. In addition, our main contributions are as follows: (1) To our best knowledge, this is the first available solution for the real-time multi-style image manipulation task – the proposed FRCNN is able to digitize over 20,000 style images in a single second. (2) Our model can handle multiple padding signals simultaneously, even in the one-shot setting (no pre-training required). FRCNN can still generate images with models comparable to previous methods that need to retrain the model for each specific padding signal (e.g., a style image with different frames of padding signals (e.g., eye patches)). (3) Using a small set of 900 multi-style images, we construct a much larger dataset of 20,000 style images, which is much larger than the dataset used in previous work. (4) Based on this dataset, we can dramatically decrease the time of training.

paragraph 1000: We thank Kiwon Choi for suggestion and participation of this work from beginning to the final stage. We also thank Kyoung Sik Joong and Cheong Sub Shin for reading off the draft and giving numerous valuable comments. This work was supported by IITP-2016-2015-02000.

©

reference

paragraph 1005: is deconvolved in the X - Z plane. It means that generically one cannot make the required pattern of PQ symmetry breaking with three fields.

We are thus led to consider even more than three fields. The simplest possibility with four fields will be:

$$H_{\text{PQ}} = \mu_1 \frac{X^{N+1}Y}{X^2Z^2} + \mu_2 \frac{X^{N+1}Z}{X^2Y^2} \quad (8.6)$$

paragraph 995: same different PQ charges from the (X, Y) fields so that any interaction between them is suppressed. Then each term will realize the correct PQ symmetry breaking pattern as in the two fields case, and we can use the two kinds of fields of different PQ charges to generate two parameters among $\{\mu, C_L, \mu'\}$ with the PQ charge assignments in Table 2. In this way, one can realize the low-level-based ENSM with one single field extended Higgs sector, but it requires such a complication of the PQ breaking sector that there must be more than three fields and a non-trivial PQ charge relation among them.

Acknowledgments

We thank Kiwon Choi for suggestion and participation of this work from beginning to the final stage. We also thank Kyoung Sik Joong and Cheong Sub Shin for reading off the draft and giving numerous valuable comments. This work was supported by IITP-2016-2015-02000.

References

- [1] E. E. Boos and G. E. Kiss, “Supersymmetry and Particle Physics,” *Phys. Rept.*, **110**, 1 (1984).
- [2] E. E. Boos and G. E. Kiss, “The Search for Supersymmetries: Probing Physics Beyond the Standard Model,” *Phys. Rept.*, **117**, 73 (1985).
- [3] J. E. Kim, “Light Fermions, Particle Physics and Cosmology,” *Phys. Rept.*, **150**, 187 (1987).
- [4] J. E. Kim and G. C. Verzini, “Axions and the Strong CP Problem,” *Rev. Mod. Phys.*, **62**, 327 (1990). [[arXiv:hep-ph/9005069](#)]
- [5] H. Minakata, H. Suzuki and T. Yanagida, “Radiative breaking of Peccei-Quinn symmetry at the intermediate mass scale,” *Phys. Lett. B*, **281**, 419 (1992).
- [6] R. D. Peccei and H. R. Quinn, “CP Conservation in the Presence of Instantons,” *Phys. Rev. Lett.*, **38**, 1450 (1977).
- [7] R. D. Peccei and H. R. Quinn, “Constitutive Imposed by CP Conservation in the Presence of Instantons,” *Phys. Rev. D*, **16**, 1796 (1977).
- [8] J. E. Kim and H. P. Nilles, “The see Problem and the Strong CP Problem,” *Phys. Lett. B*, **198**, 159 (1988).
- [9] D. M. Lytle and E. D. Stewart, “Densest Inference and the moduli problem,” *Phys. Rev. D*, **58**, 1750 (1998). [[hep-ph/9502005](#)]
- [10] K. Choi, E. J. Choi and J. E. Kim, “Cosmological implications of radiatively generated axion scale,” *Phys. Lett. B*, **665**, 269 (2008). [[hep-ph/0609022](#)]
- [11] M. Papucci, J. T. Ruderman and A. Weiler, “Natural SUSY Eardens,” *JHEP*, **1209**, 031 (2012). [[arXiv:1110.0986](#)]

paragraph 1000: *Fast R-CNN* for image manipulation scheme for the multi-style image manipulation, that performs better than the previous approaches with a low computational complexity of $O(1)$ (eq. 8.1). The computation time can be reduced by the proposed algorithm. The proposed algorithm proves that uses a prior condition on the divergence in the form of a subgraph, and correctly handles the PP divergence in the PP grid, while reducing noise. The resulting images are sharper and more realistic. The proposed algorithm is efficient and is capable to reduce noise and achieve high accuracy. It is shown to perform significantly better than known denoising techniques, either operating in the spatial or frequency domains. In addition, we showed that a PP grid can be used to handle the noise that is not present in the original image. The proposed PP algorithm produces superior results than state-of-the-art algorithms, with a lower computational complexity and in less time. The method can also be adapted to medical images and for brain segmentation in CT scans. We hope this work will pave the way for a CT scanner that uses the PP paradigm for producing clinical images that are taken with significantly less radiation dosage.

paragraph 995: *Faster R-CNN* for image manipulation scheme for the multi-style image manipulation, that performs better than the previous approaches with a low computational complexity of $O(1)$ (eq. 8.1). The computation time can be reduced by the proposed algorithm. The proposed algorithm proves that uses a prior condition on the divergence in the form of a subgraph, and correctly handles the PP divergence in the PP grid, while reducing noise. The resulting images are sharper and more realistic. The proposed algorithm is efficient and is capable to reduce noise and achieve high accuracy. It is shown to perform significantly better than known denoising techniques, either operating in the spatial or frequency domains. In addition, we showed that a PP grid can be used to handle the noise that is not present in the original image. The proposed PP algorithm produces superior results than state-of-the-art algorithms, with a lower computational complexity and in less time. The method can also be adapted to medical images and for brain segmentation in CT scans. We hope this work will pave the way for a CT scanner that uses the PP paradigm for producing clinical images that are taken with significantly less radiation dosage.



Figure 5: Example output of Faster R-CNN on the test set

paragraph 1005: *Fast R-CNN* for image manipulation scheme for the multi-style image manipulation, that performs better than the previous approaches with a low computational complexity of $O(1)$ (eq. 8.1). The computation time can be reduced by the proposed algorithm. The proposed algorithm proves that uses a prior condition on the divergence in the form of a subgraph, and correctly handles the PP divergence in the PP grid, while reducing noise. The resulting images are sharper and more realistic. The proposed algorithm is efficient and is capable to reduce noise and achieve high accuracy. It is shown to perform significantly better than known denoising techniques, either operating in the spatial or frequency domains. In addition, we showed that a PP grid can be used to handle the noise that is not present in the original image. The proposed PP algorithm produces superior results than state-of-the-art algorithms, with a lower computational complexity and in less time. The method can also be adapted to medical images and for brain segmentation in CT scans. We hope this work will pave the way for a CT scanner that uses the PP paradigm for producing clinical images that are taken with significantly less radiation dosage.

paragraph 995: *Faster R-CNN* for image manipulation scheme for the multi-style image manipulation, that performs better than the previous approaches with a low computational complexity of $O(1)$ (eq. 8.1). The computation time can be reduced by the proposed algorithm. The proposed algorithm proves that uses a prior condition on the divergence in the form of a subgraph, and correctly handles the PP divergence in the PP grid, while reducing noise. The resulting images are sharper and more realistic. The proposed algorithm is efficient and is capable to reduce noise and achieve high accuracy. It is shown to perform significantly better than known denoising techniques, either operating in the spatial or frequency domains. In addition, we showed that a PP grid can be used to handle the noise that is not present in the original image. The proposed PP algorithm produces superior results than state-of-the-art algorithms, with a lower computational complexity and in less time. The method can also be adapted to medical images and for brain segmentation in CT scans. We hope this work will pave the way for a CT scanner that uses the PP paradigm for producing clinical images that are taken with significantly less radiation dosage.

- [1] S. Bruck, *Computer Graphics: Principles, Design, Artifacts, and Recent Advances*. Birkhäuser, Wien, 2003.
- [2] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*. IEEE press, 1989.
- [3] L. A. Feldman and L. C. Davis, and W. K. Kerr, “Practical cone-beam algorithm,” *Journal of the Optical Society of America A*, vol. 1, no. 6, p. 612 (1984).
- [4] A. Arora, R. R. Coifman, D. L. Donoho, M. Elad, and J. Ward, “A fast dual-tree complex wavelet transform for data in a Cartesian grid which is nearly complete, algebraically robust, geometrically tight, and invertible,” *SIAM Journal on Scientific Computing*, 2005.
- [5] V. Arachchi, R. Colomo, and D. Donoho, “A framework for discrete integral transformations I: the pentagonal Fourier transform,” *SIAM Journal on Scientific Computing*, vol. 30, no. 2, pp. 760–780, 2008.
- [6] A. Arora, R. R. Coifman, D. L. Donoho, T. Minkowski, and I. Saban, “A framework for discrete integral transformations II: the 2D dual-tree Radon transform,” *SIAM Journal on Scientific Computing*, vol. 30, no. 3, pp. 1830–1850, 2008.
- [7] P. A. Paliogianni, A. G. Lindqvist, “Sampling the 2D Radon transform: IEEE Transactions on Acoustics, Speech, and Signal Processing”, *vol. 28*, no. 3, pp. 994–1002, 1980.
- [8] V. C. Elkjaer, *Sampling Theory: General Fundamentals*. Springer, 2002.
- [9] A. Beck and M. Teboulle, “Fast gradient-based algorithms for constrained total variation denoising and deblurring problems,” *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2432–2442, 2009.
- [10] D. P. Palomar and C. E. Kokkila, *Convex Optimization in Signal Processing and Communications*, Cambridge university press, 2008.
- [11] E. Lee, R. P. Feddema, C. V. Jones, C. Sabuncu, G. E. Boos, R. K. Wright, D. C. Gitter, G. E. Kiss, and J. M. Cline, “A fast dual-tree complex wavelet transform and image enhancement in helical CT imaging through rapidly-spiraled tomography,” *Journal of Biomedical Optics*, vol. 16, no. 2, pp. 221–227, 2008.
- [12] E. Lee, R. P. Feddema, C. V. Jones, and J. M. Cline, “Low-dose x-ray phase contrast and derivative CT using rapidly-spiraled tomography,” *Physics in Medicine and Biology*, vol. 55, no. 18, p. 5403–5408, 2010.
- [13] Y. Mai, R. P. Feddema, S. J. Oliver, and J. M. Cline, “Derivative tomography using a regularized Bayesian approach: Reconstruction and Optimization,” *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1259–1269, 2010.
- [14] D. M. Matthews, S. Baladandi, and P. O’Gallagher, “Fast iterative Low-dose X-ray CT Reconstruction through Sparse-Rank MIP Modeling,” *arXiv preprint*, vol. arXiv:1602.01105, pp. 1–10, 2016.
- [15] H. Kudo, M. Ueda, M. Ueda, and J. Yamada, “A novel approach to image reconstruction using sparse representation,” *arXiv preprint*, vol. arXiv:2006.15564, 2020.
- [16] Alexander Achille and Stefano Soatto, “Information dropout: Learning optimal representations through noisy computation,” *ArXiv e-prints*, 2016.
- [17] Alexander Achille and Stefano Soatto, “On the Emergence of Invariance and Disentanglement in Deep Representations,” *arXiv e-prints*, June 2017.
- [18] Alessandro Achille and Stefano Soatto, “Deep information for learning and control,” *SIAM Review*, vol. 59, pp. 1037–1061, 2017.
- [19] Alessandro Achille and Stefano Soatto, “Deep information for learning and control,” *SIAM Review*, vol. 59, pp. 1037–1061, 2017.
- [20] Alessandro Achille and Stefano Soatto, “Deep variational information bottleneck,” *arXiv preprint*, arXiv:1807.08030, 2018.
- [21] Alessandro Achille and Stefano Soatto, “Balanced approximation of stochastic systems,” *SIAM journal on matrix analysis and applications*, 1993.
- [22] Bagdasaryan, Andreev, and Vardanyan, “Sufficiency and optimality conditions for the theory of representations and implementing efficient algorithms to solve inverse problems,” *ArXiv e-prints*, 2017.
- [23] Bagdasaryan, Andreev, and Vardanyan, “On the Emergence of Invariance and Disentanglement in Deep Representations,” *arXiv e-prints*, 2017.
- [24] Bagdasaryan, Andreev, and Vardanyan, “Deep variational information bottleneck,” *arXiv preprint*, arXiv:1807.08030, 2018.
- [25] Bagdasaryan, Andreev, and Vardanyan, “Balanced approximation of stochastic systems,” *SIAM journal on matrix analysis and applications*, 1993.
- [26] Bagdasaryan, Andreev, and Vardanyan, “Sufficiency and optimality conditions for the theory of representations and implementing efficient algorithms to solve inverse problems,” *ArXiv e-prints*, 2017.
- [27] V. Bagdasaryan and T. E. Hartung, *Ranking and Data Association*. Academic Press, 1987.
- [28] Bagdasaryan, Andreev, and V. Hartung, “Balanced diagrammatic reduction of representations with categorical products,” *Annals of Mathematics*, pp. 675–693, 2016.
- [29] Thomas M. Cover and Joy A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [30] Laurent Dinh, Razvan Pascanu, Samy Bengio, and Yoshua Bengio, “Sharp minima can generalize for deep nets,” *arXiv preprint*, arXiv:1705.04643, 2017.
- [31] Shlomo Shamir and Vladislav Kostin, “Learning to act by predicting the future,” *arXiv preprint*, arXiv:1707.00001, 2017.
- [32] Ray Liu, Michael Mathieu, and Yoshua Bengio, “Principled option learning in markov decision processes,” *arXiv preprint*, arXiv:1609.07524, 2016.
- [33] Ray Liu and Yoshua Bengio, “Minimum information lag control in reinforcement controllers,” *Neurocomputing*, pp. 5603–5608, 2018.
- [34] Yann LeCun, Léon Bottou, Yoshua Bengio, and Yves Grandvalet, “Classifying digits with a locally linear discriminant,” *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pp. 249–256, 2010.
- [35] Bagdasaryan and Vardanyan, “Balanced approximation of stochastic systems,” *SIAM Computational Mathematics*, 1993.
- [36] Mata-Montes, M. Chen, Yen-Duan, John Schreiber, Philip De Souza, and Peter Attia, “Balanced diagrammatic renormalization exploration,” *Advances in Neural Information Processing Systems (NIPS)*, 2016.
- [37] Andrew H. Trivedi, “Stochastic processes and filtering theory,” *Courier Corporation*, 2007.
- [38] Max J. Newbold, “An extension of the principal component theorem,” *In Mathematical Proceedings of the Cambridge Philosophical Society*, volume 36, pages 383–395. Cambridge University Press, 1940.
- [39] Michael E. Kalman, “A new approach to linear filtering and prediction problems,” *ASME Journal of Dynamic Systems, Measurement, and Control*, 1960.
- [40] David Kriegsmann and James B. Keller, “A method for stochastic optimization,” *arXiv preprint*, arXiv:0311.2000, 2003.

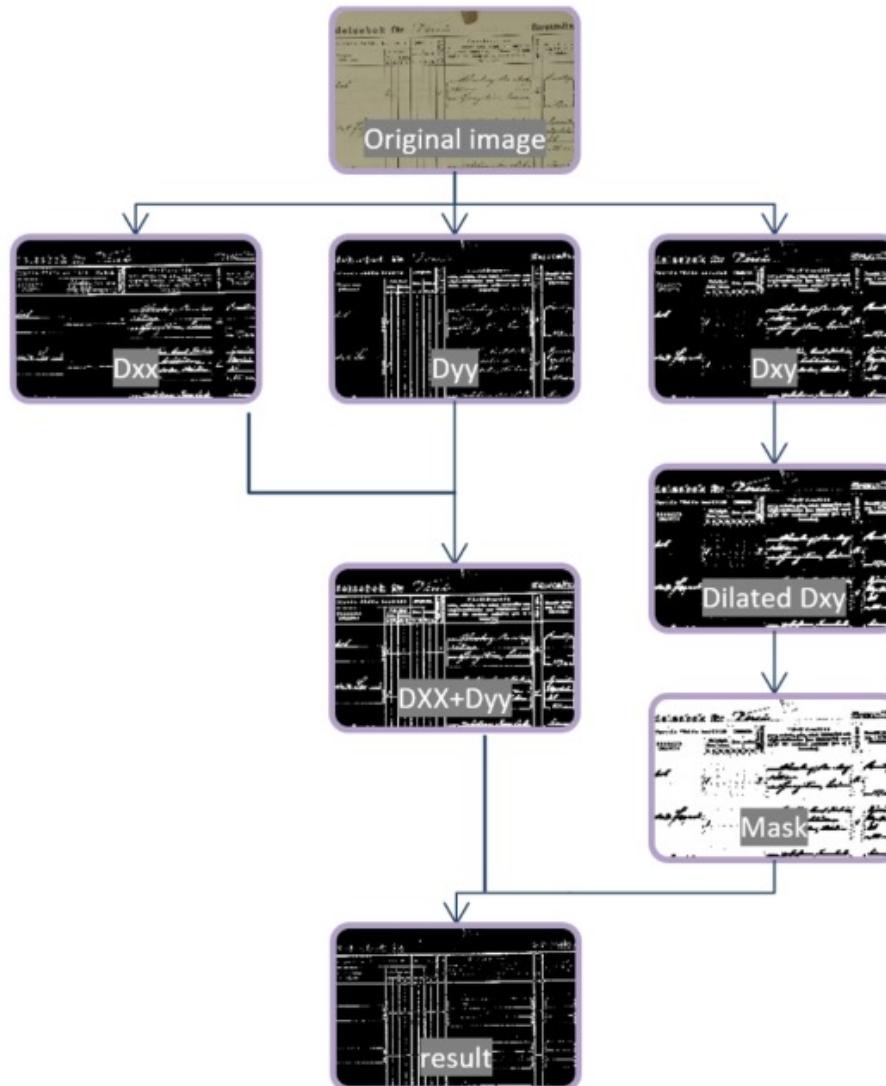


Fig. 6. Process flow of mask image production using 2D Hessian filters.

Liang et al. (2021), Comparative Study of Layout Analysis of Tabulated Historical Documents, DOI: [10.1016/j.bdr.2021.100195](https://doi.org/10.1016/j.bdr.2021.100195)



Fig. 5. Example of a document image with a skew of 5 degrees.

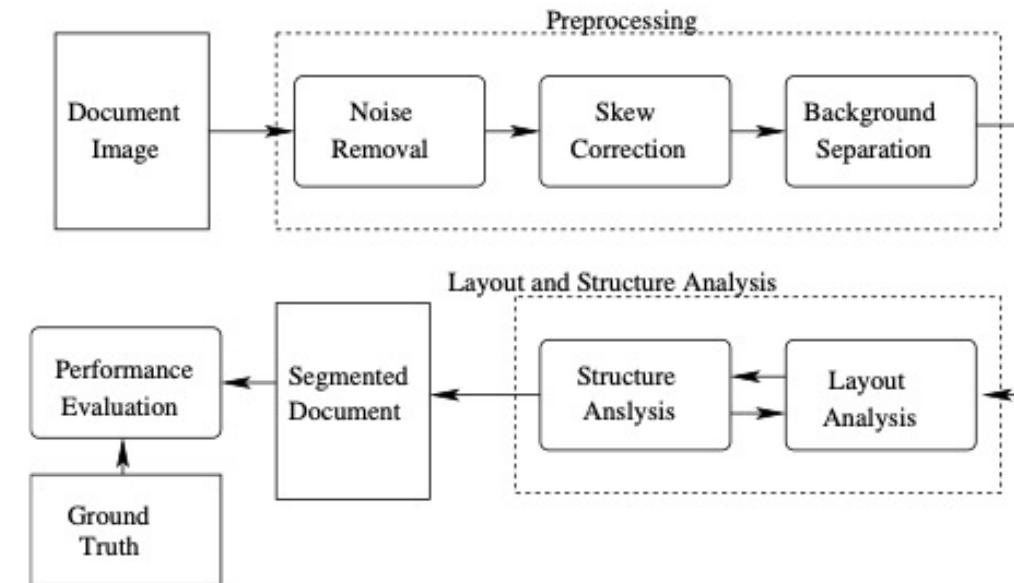


Fig. 2. Schematic diagram of a document layout and structure analysis system.

GNHK: A dataset for English handwriting in the wild

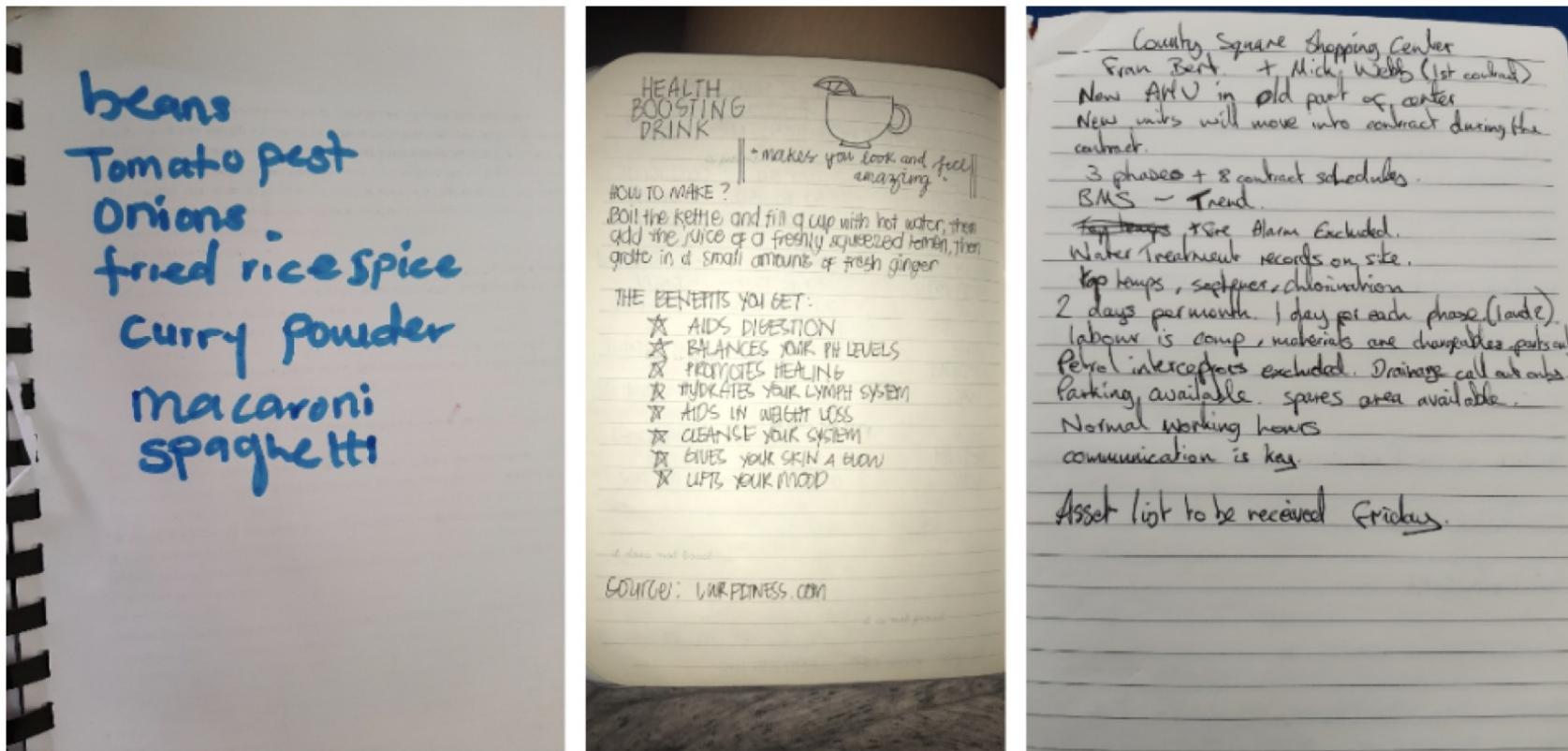
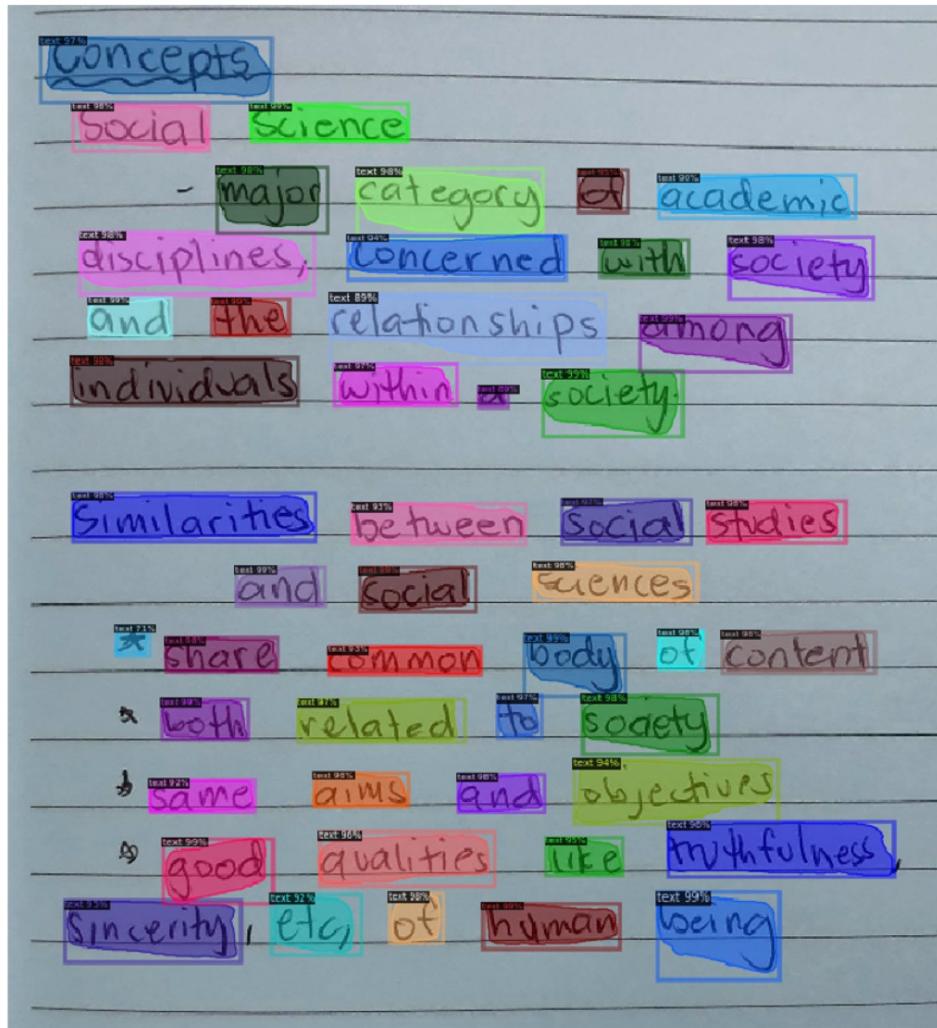
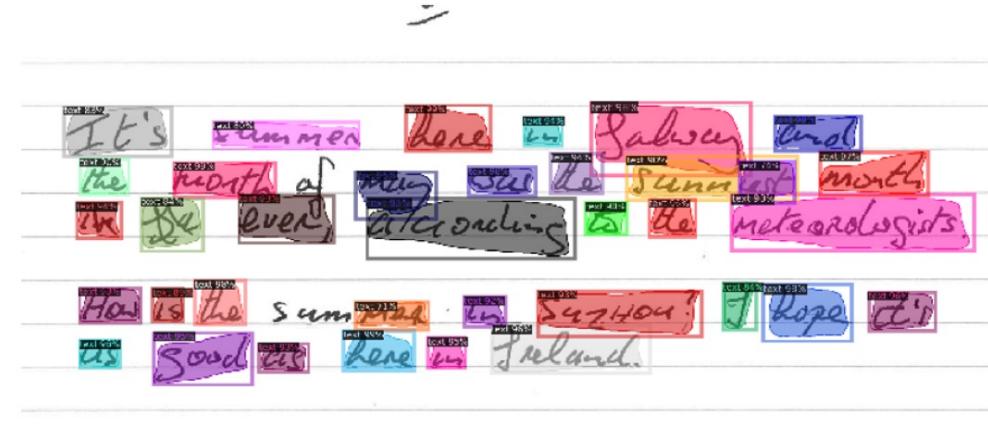


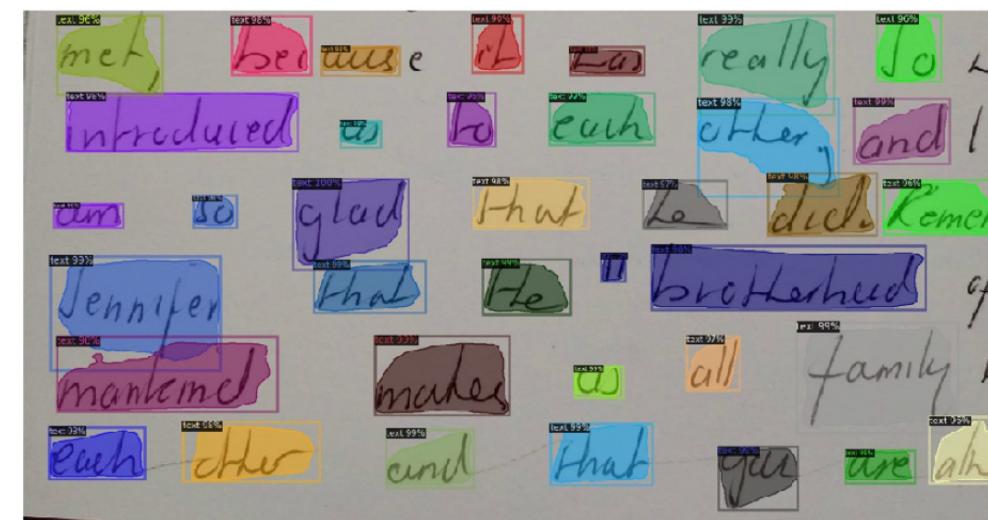
Fig. 1: Examples of the handwriting images captured under unconstrained settings.



(a)



(b)



(c)

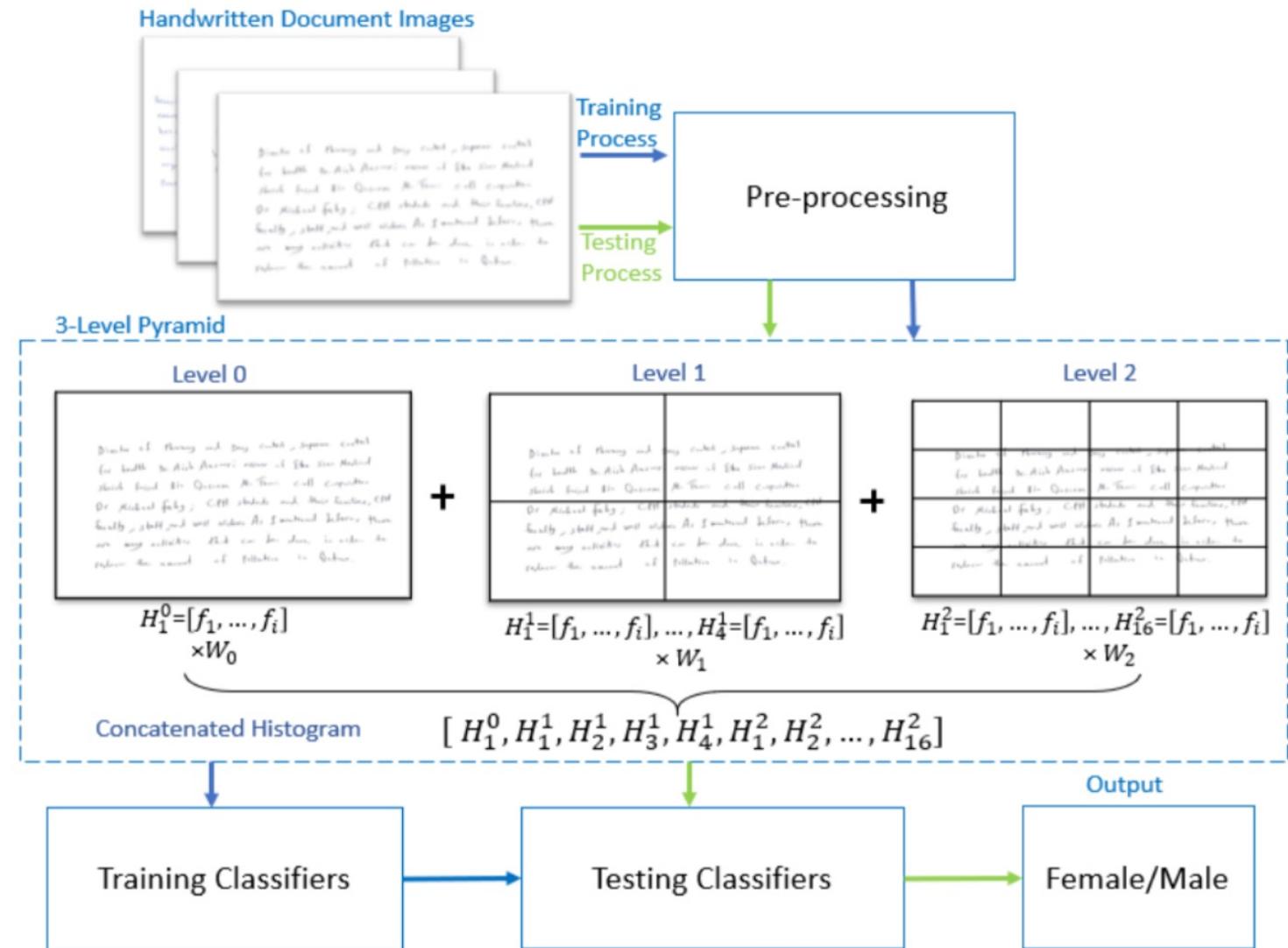
Fig. 8: Examples of the Mask R-CNN results.

Gender Detection

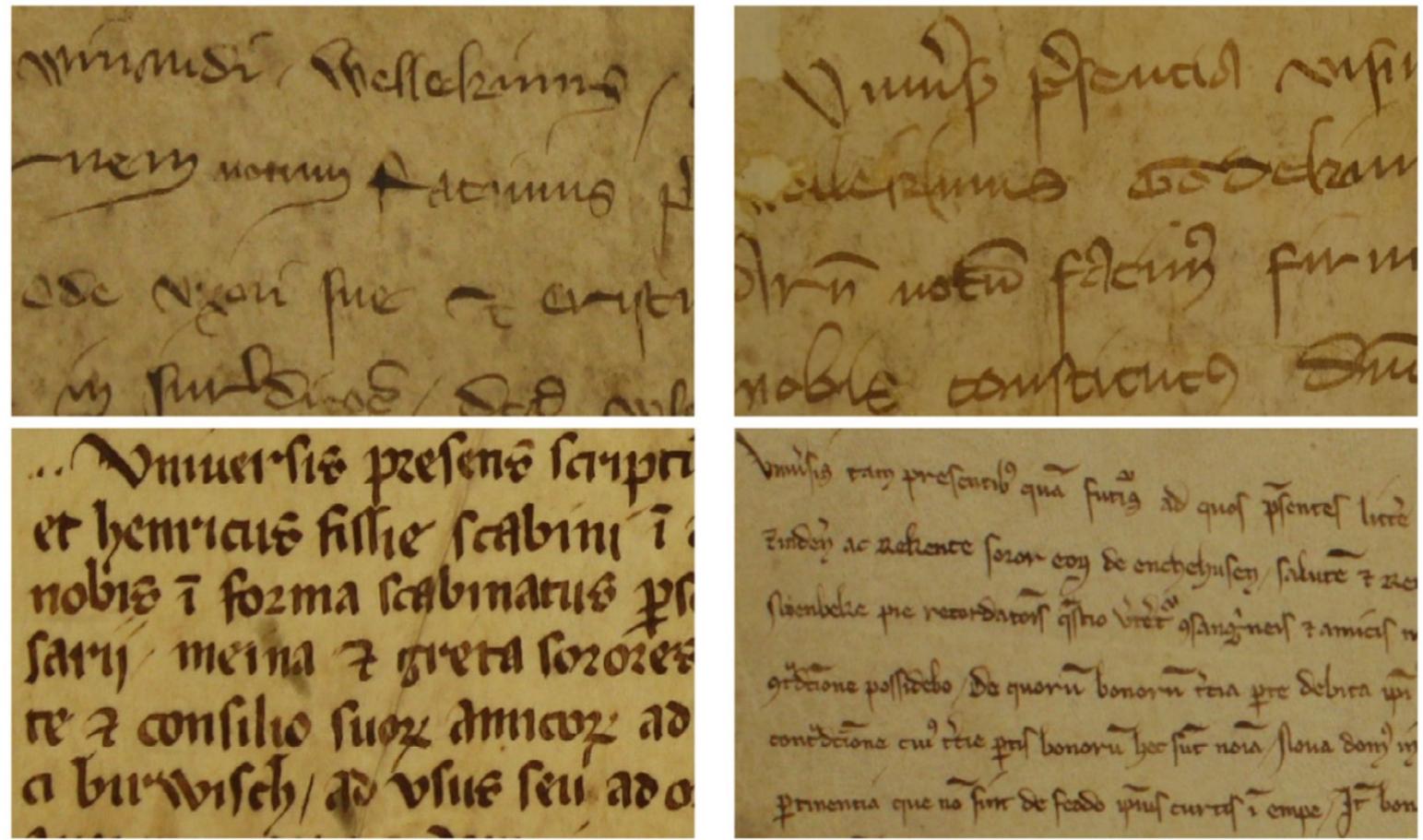
Local Feature – SIFT

Visual Vocabulary – K-Means

Classifier – SVM, AdaBoost



Historical manuscript dating



Sheng et al., (2016). Historical manuscript dating based on temporal pattern codebook
<https://www.sciencedirect.com/science/article/pii/S1077314216301163/pdf>

Historical manuscript dating

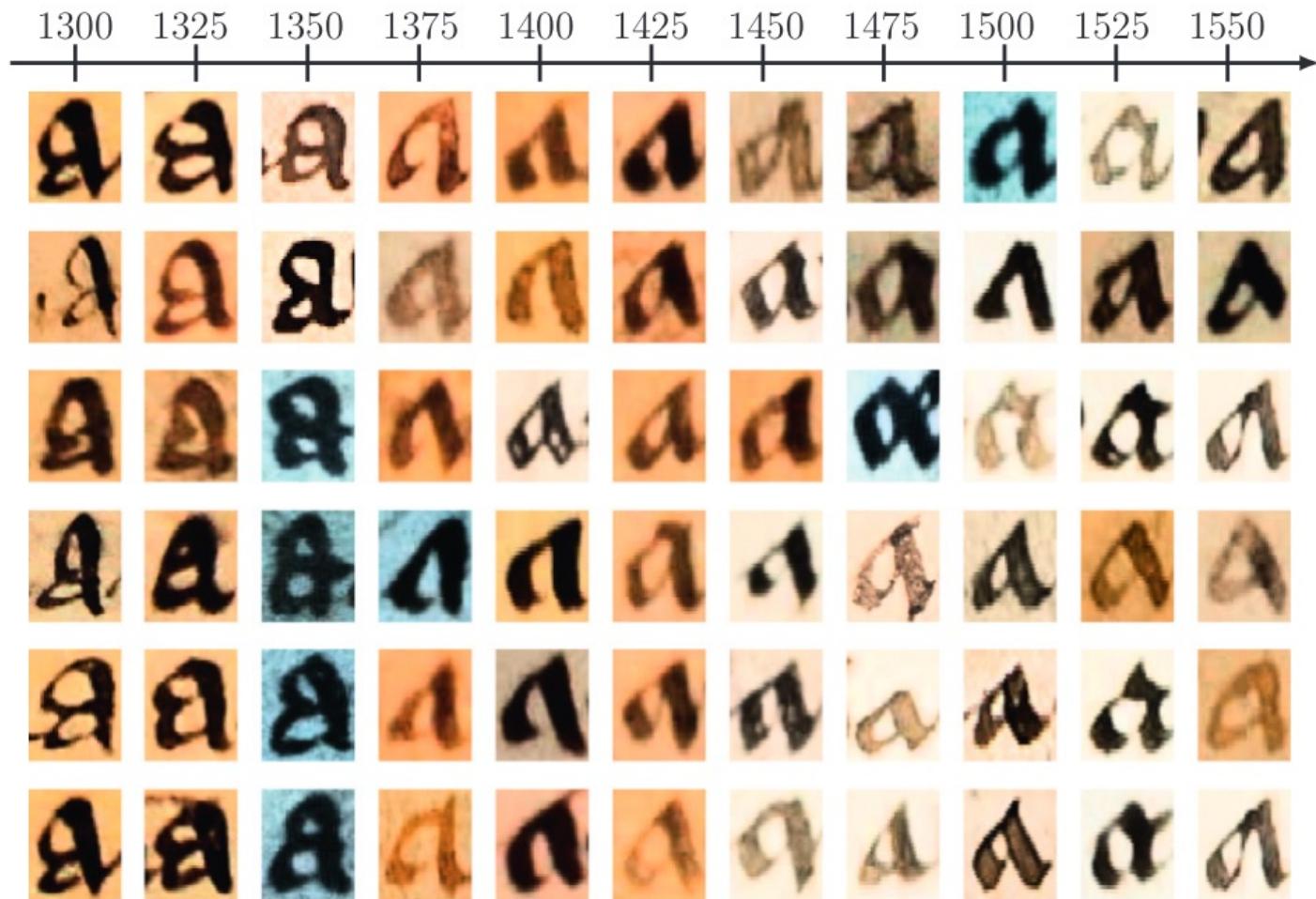


Fig. 1. Six randomly selected instances of the character 'a' in the 11 key years in our MPS data set.

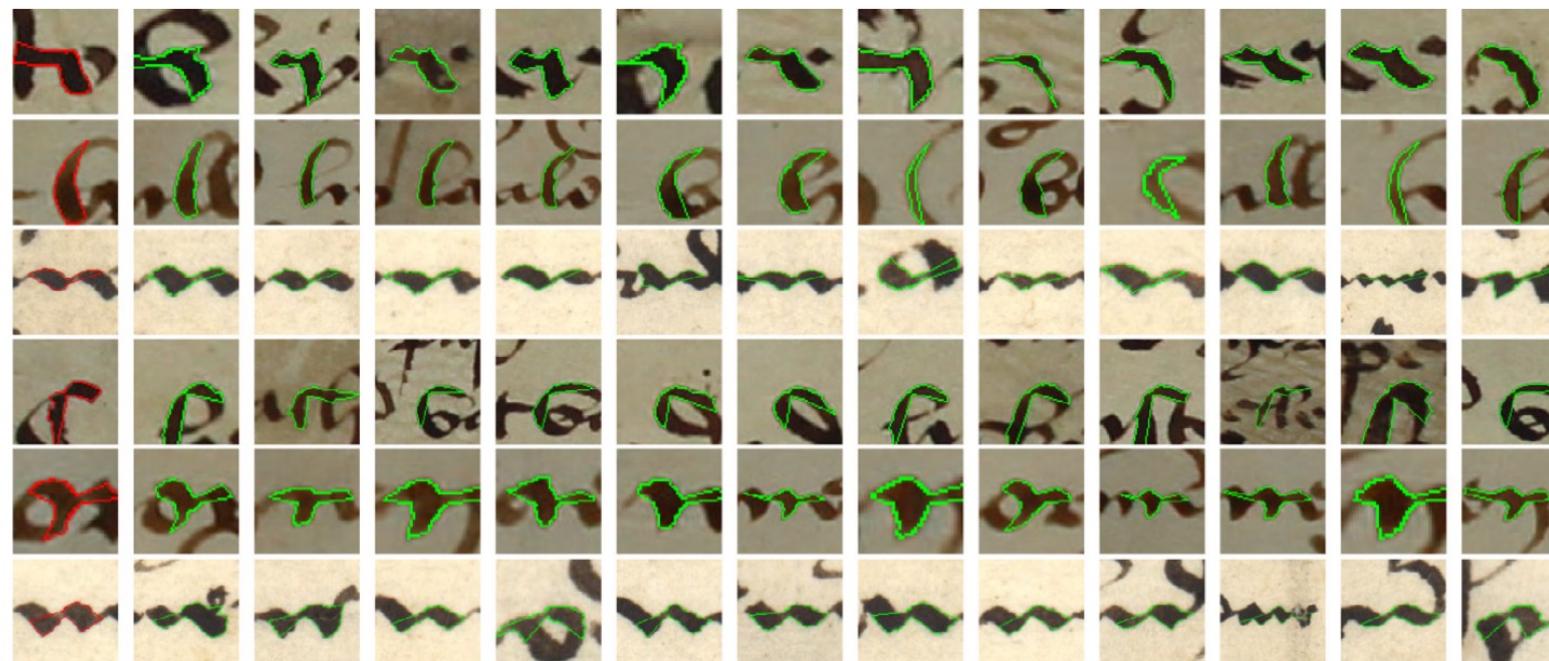
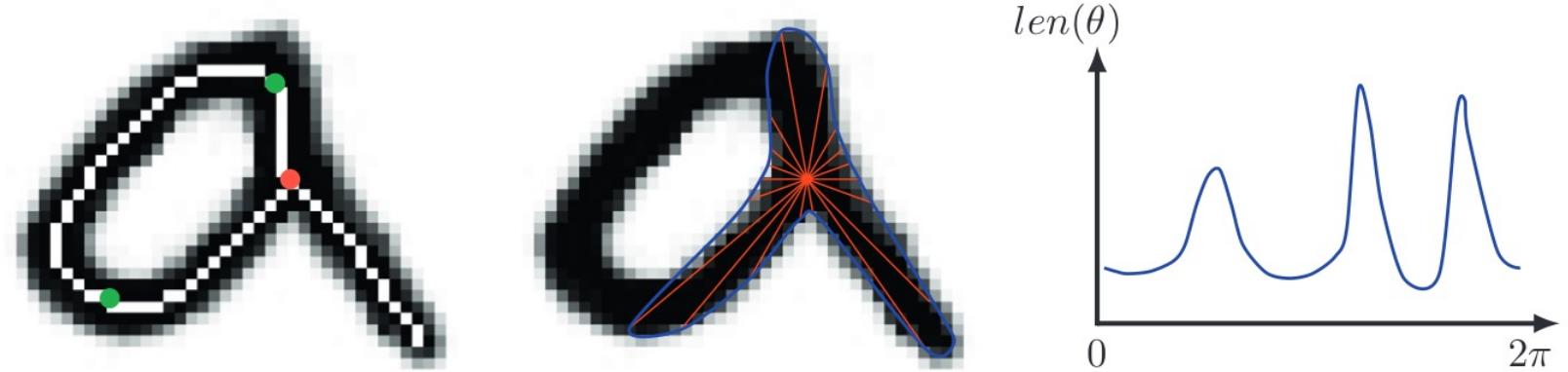


Fig. 5. Each row shows the first 12 instances in a hit list of the query patterns of the first column. Note that each patch is normalized into a fixed size.

Thank you for your attention