Ben-Gurion University of the Negev
Department of Computer Science

Deep Learning for Historical Document Image Analysis

למידה עמוקה לניתוח תמונת מסמכים היסטורית

Thesis submitted in partial fulfillment
of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Berat Kurar Barakat

Submitted to the Senate of
Ben-Gurion University of the Negev

July 2021

Be'er-Sheva

Ben-Gurion University of the Negev
Department of Computer Science

Deep Learning for Historical Document Image Analysis

למידה עמוקה לניתוח תמונת מסמכים היסטורית

Thesis submitted in partial fulfillment
of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Berat Kurar Barakat

Submitted to the Senate of
Ben-Gurion University of the Negev

Approved by the advisor Prof. Jihad El-Sana:
Approved by the Dean of the Kreitman School of Advanced Graduate Studies:

July 2021

Be'er-Sheva

This thesis was carried out under the supervision of

Professor Jihad El-Sana

Department of Computer Science
Faculty of Natural Sciences

I the undersigned, Berat Kurar Barakat, hereby declare that

I have written this Thesis by myself, except for the help and guidance offered by my Thesis Advisors.

The scientific materials included in this Thesis are products of my own research, culled from the period during which I was a research student.

This Thesis incorporates research materials produced in cooperation with others, excluding the technical help commonly received during experimental work. Therefore, I am attaching another affidavit stating the contributions made by myself and the other participants in this research, which has been approved by them and submitted with their approval.

Date: _____ Name: _____ Signature: _____

Acknowledgements

I would like to thank my advisor, Professor Jihad El-Sana, for his guidance during my thesis work.

I would also like to thank my wonderful parents, Gunes and Huseyin, for their endless love.

I am grateful to my dear husband Emaduddien for his continuous support and care.

Most importantly, a special feeling of gratitude goes to my lovely sons, Fatih and Hamza.

Abstract

Much recent research have been devoted to historical document image analysis, as automatic understanding of documents from images enables scholars to reach the relevant data easily. Conventional methods require domain experts to carefully design feature extractors, and they are limited in their ability to process raw input images. Deep learning models are composed of multiple layers that are able to learn task specific representations of data, with multiple levels of abstraction directly from raw input images. These models have demonstrated remarkable success in computer vision, hence there has been a trend of using deep learning for historical document image analysis.

This thesis is a set of articles that propose several deep learning models to formulate a variety of document image analysis tasks. First, we use a fully convolutional network to segment document images into homogeneous regions. Next, we formulate text line detection as dense prediction of blob lines that strike through text lines. Text line detection is a loose definition that provides only the spatial locations of text lines. Hence, we extend text line detection to text line extraction by minimizing an energy function that is guided by the detected blob lines. However, training deep learning models relies on the availability of exhaustive data annotations, limiting their performance in rare cases such as curved lines and scalability in new application sets. Therefore, we investigate using Laplacian of multi-oriented anisotropic Gaussians to segment multiple oriented and curved text lines. Additionally, we introduce two unsupervised deep learning models for text line segmentation, which require no manual annotation. We also propose a learning free method that can adapt to heterogeneity by automatic scale selection. Finally, we introduce the Pinkas dataset that is the first dataset in medieval handwritten

Hebrew and fully labeled at word, line and page level by a Hebrew paleographer. We perform baseline experiments with three methods for word spotting, to serve as a benchmark for future studies and to mark a transformation in research for analyzing Hebrew handwriting.

Contents

1 Introduction

Libraries provide access to digital copies of historical documents for promoting worldwide accessibility to cultural heritage, while preserving their physical copies from further deterioration caused by recurrent usage. Digital document images are not easy to explore in their raw form and need to be transcribed further into machine readable text. Certainly, manual transcription for huge amounts of document images is not feasible in a reasonable time. Therefore, there is a significant need for reliable historical document image processing algorithms. This thesis focuses on page segmentation, text line segmentation and word spotting algorithms for historical document images.

1.1 Research problems

Typically, document image analysis first segments a document page image into text and non-text regions. Then, text regions are segmented into text lines which are segmented into words. Finally, the recognition is performed on the segmented text lines or segmented words because it can be difficult to segment a handwritten text into characters. This thesis investigates how to model page segmentation, text line segmentation and word spotting as a problem solvable using deep learning algorithms.

1.1.1 Page segmentation

Page segmentation is the process of detecting the homogeneous regions of document images. The recognition process targets the text regions, therefore the de-

tected regions need to be labeled as either text regions or non-text regions. The detected regions are represented by bounding polygons that cover the foreground and background pixels together. The handwritten documents pose problems such as multiply oriented and curved text regions, arbitrarily shaped complex layout, multiple font types and sizes in a single document image.

1.1.2 Text line segmentation

Text line segmentation is the process of detecting the regions that contain a single text line. Detected regions can be represented in coarse or fine forms. The coarse form is called text line detection and it represents a text line by its spatial location. This representation can be expressed by baselines or blob lines. A baseline is a polyline that strikes through the bottom part of the main body of the characters in the detected text line. A blob line is a pixel blob that masks the main body regions of the characters in the detected text line. The fine form is called text line extraction and it represents a text line by labeling the foreground pixels that belong to the extracted text line. This representation can be expressed by pixel labeling or bounding polygons. Pixel labeling labels all the foreground pixels of the text line with the same label. Bounding polygon is a polygon that covers all the foreground pixels of a text line together with some background pixels among them. The segmentation of handwritten text lines is particularly difficult because of the multiply oriented and skewed marginal notes, touching and overlapping characters, multiple font types and sizes, crowded diacritics, and cramped text lines.

1.1.3 Word spotting

Word spotting is the task of matching a query word image with all the word images in the documents, and retrieve the similar word images to the given query image. Conventional methods for recognizing a given text suffer from degradation due to stains and inks, multiple font types and sizes, and cursive writing, when used with historical handwritten documents. Therefore, word spotting detects the locations on document images which contain an instance of the queried word without ex-

plicitly recognizing the word. The query word can be represented by means of an example image or a string of characters which is relatively easier for the user.

1.2 Literature survey

Humanities scholars have long desired of computerized tools that transcribe historical document images and make them easy to explore. In the early days of document image analysis, the researchers tackled simple problems that can be described by handcrafted features. Simple problems arise from rectangular layouts, horizontal or slightly skewed spacious text lines with bare components, almost homogeneous font types and font sizes, and fixed size vocabulary of words. The true challenge to document image analysis proved to be solving the problems that are hard to describe by mathematical rules. Hard problems arise from complex layout, multiple oriented and curved, cramped text lines crowded by satellite components, heterogeneous font types and font sizes, and variable vocabulary of words.

Since then, the AlexNet [31] achieved a top-5 test error rate almost halved the error rate of the second best entry that uses hand crafted features, numerous deep learning based methods have been proposed for document image analysis tasks. The key difference of a deep learning model is that it learns not only the mapping function from input to output but also the features. The learned features eliminate the challenge of manually describing mathematical rules for hard problems in document image analysis.

Rectangular layout page segmentation problems can be solved using Run Length Smearing Algorithm (RLSA) [66], projection profiles [1] or white tiles, [5] whereas complex layout document images can be segmented by classifying keypoint features [13, 17, 22] or clustering texture features [7, 47]. Recently complex layout page segmentation has achieved state of the art results using unsupervised Convolutional Neural Network (CNN) [16, 21, 65], supervised CNN [15, 42] and supervised Fully Convolutional Network (FCN) [49, 67].

Horizontal text line segmentation problems are solved using projection profiles based methods, which were first applied to documents with horizontal text lines

[26, 45], and subsequently adapted to documents with slightly skewed [6, 9] and multi-skewed text lines [50]. Gray scale document images can be segmented into text lines using seam carving methods [2, 8, 57, 58]. Multiply oriented, curved and cramped text lines are widely solved by clustering [25, 51] or enhancing the text line structure [14, 18, 25, 51].

Deep learning of handwritten text line segmentation has a recent history starting from the Recurrent Neural Network (RNN) work of Moyyset et al. [48]. The sequential characteristic of RNN limits the input to the form of horizontal text lines. Therefore, researchers have started focusing on layout agnostic formulation of the text line segmentation. The potential formulation has to define which pixels belong to a text line and predict whether a pixel is a text line pixel or not. Defining a text line with its raw foreground pixels would yield semantic segmentation to be paradoxically ineffective, because the output will not discriminate a text line from the others. This inherent limitation has caused to define text lines as a single connected component. This component can be either a blob line [35, 37, 46, 54, 64] strikes through the main body area of the characters that belong to a text line or a baseline [24] passes through the bottom part of the main body of the characters that belong to a text line.

Recognition free word retrieval was initially proposed by Manmatha et al. [44] and is called word spotting. It avoids recognition of words, but matches each word image against all the other word images. Matching techniques fall into two categories: pixel by pixel matching and feature based matching. Pixel by pixel matching compares two images pixel by pixel using XOR [28, 43], Sum of Squared Differences (SSD) [43] or Euclidean Distance Mapping (EDM) [28, 43]. Feature based matching extracts image features and compares two images using Scott and Longuet Higgins (SLH) [59] algorithm [28, 43], Shape Context (SC) algorithm [11, 52], Dynamic Time Wrapping (DTW) [52] or corner feature correspondence algorithm (CORR) [56]. In general, feature based matching, and in specific DTW, performs best among this set of algorithms surveyed in [53]. Histogram of Gradients (HOG) and Scale Invariant Feature Transform (SIFT) features have also been used for word spotting [55, 63]. These gradient based features can represent the directions of the strokes and are not constrained to single writer collections. The

state of the art results have been achieved by using a CNN as feature extractor [30, 60, 62].

Deep learning models can inherently handle the hard problems arising from complexity and heterogeneity of the document images. However, they require a vast amount of labeling effort which consumes time not less than designing features manually. Intuitively, labeling effort is favorable over manual feature design because the former can be accomplished by human recognition skills, whereas the latter requires further mathematical skills. We identify the recent works related to page segmentation, text line segmentation and word spotting in the following subsections.

1.3 Thesis survey

This doctoral dissertation aims to analyse historical handwritten document images using deep learning algorithms. At first, we formulated page segmentation [40] and text line detection [37] as dense prediction problems. Observing that supervised learning is not successful at curved text line detection due to annotated data scarcity, we solved it by a learning-free algorithm [34]. In the meantime, we realized that document image analysis problems are usually benchmarked on Latin or Arabic handwritten datasets, and so we introduced the first Hebrew handwritten dataset [41] for page, line and word segmentation tasks. We further formulated an energy minimization function [35] that can extract text lines by the guidance of detected text lines with any orientation, font type and font size. We experienced the enormous effort of data annotation and proposed a learning-free method [33] and two unsupervised deep learning methods [36, 39] for the text line segmentation problem.

1.4 Methods

Fundamental to our work in this thesis are CNN and FCN besides scale space of Laplacian of anisotropic Gaussians and energy minimization. The thesis intro-

duces several models built from CNN and FCN for solving page segmentation, text line segmentation and word spotting problems. Scale space of Laplacian of anisotropic Gaussians is used as an alternative solution for text line segmentation when labelled data is scarce. Both models for text line segmentation, learning-free or learning-based, are accompanied by energy minimization in the post processing phase. This section provides an overview of these methods used in the articles, and our models built from these methods are to be presented in the articles with details.

1.4.1 Convolutional neural network

A neural network is a series of an input vector, hidden layers, and an output vector that represents the class scores for the given input. Each hidden layer is a set of independent neurons that have learnable weights and biases. Each neuron is fully connected to all the neurons in the previous layer. This fully connected structure does not scale well to image data, as a flattened image would lead to a huge number of parameters that would cause overfitting. Hence, a convolutional neural network assumes that the input is an image and arranges the neurons in 3D volume: height, width and depth. Every cell in the 3D volume is an output of a neuron that looks at a small region, namely receptive field, in the input volume and shares weights with all neurons across the height and width. The weight sharing is intuited from the fact that a feature useful in one part of the image is probably useful in another part of it.

A CNN is made up of input, convolutional, pooling and fully connected layers. Every layer transforms its input 3D volume to an output 3D volume. Input layer consists of raw image pixels. Convolutional layer without neuron analogy consists of a set of filters whose weights correspond to the input weights of neurons. Convolutional layer convolves each filter across the height and the width of the input volume and applies element wise non-linearity such as Sigmoid, ReLU and Tanh. Each filter produces a 2D activation map that is stacked along the depth dimension to produce the output 3D volume. Pooling layer downsamples the height and width dimensions to reduce the number of weights in the network, and to al-

low small translational invariances. Fully connected layer neurons are connected to all the neurons in its input volume and compute the class scores. The weights in the convolutional and fully connected layers are trained with gradient descent to converge the class scores to the labels in the training set. Eventually, early convolutional layer filters learn to extract simple features, and final convolutional layer filters learn to hierarchically extract complex features that are combinations of simple features.

1.4.2 Fully convolutional network

Convolutional networks take fixed sized inputs and produce non-spatial outputs because of the fully connected layers inherent to the class scores. Fully convolutional networks throw away the fully connected layers, take inputs of any size and produce spatial output maps. The spatial output maps enable fully convolutional networks to make dense pixel predictions for semantic segmentation given the ground truth at every output pixel. Typical convolutional networks reduce the spatial input dimensions by downsampling to keep number of parameters reasonable. The downsampling enables filters to see coarser information with a larger receptive field. Additionally, FCN upsamples coarse outputs to dense pixels. Upsampling is done by transpose convolution that is to apply a convolution filter with a stride equal to $\frac{1}{f}$, for upsampling by a factor f.

1.4.3 Scale space of Laplacian of anisotropic Gaussians

Given a continuous image $I : \mathbb{R}^2 \to \mathbb{R}$, its scale-space of anisotropic Gaussians $L : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ is defined by following convolution:

$$L(x,y;\sigma_x,\sigma_y) = I(x,y) * g(x,y;\sigma_x,\sigma_y), \tag{1.1}$$

where $g : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ denotes the anisotropic Gaussian kernel

$$g(x,y;\sigma_x,\sigma_y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{x^2}{2\sigma_x{}^2}+\frac{y^2}{2\sigma_y{}^2}\right)}. \tag{1.2}$$

In this representation $*$ is the convolution operator, σ_x is the scale parameter in horizontal direction, and σ_y is the scale parameter in vertical direction.

Scale space of Laplacian of anisotropic Gaussians is computed by differentiating the scale-space of anisotropic Gaussians with respect to x and y two times

$$L_{x^2y^2}(x,y;\sigma_x,\sigma_y) = \partial_{x^2y^2} L(x,y;\sigma_x,\sigma_y), \tag{1.3}$$

or, equivalently, by convolving the image with Laplacian of anisotropic Gaussians

$$L_{x^2y^2}(x,y;\sigma_x,\sigma_y) = I(x,y) * [g_{x^2y^2}(x,y;\sigma_x,\sigma_y)]. \tag{1.4}$$

Given the anisotropic Gaussian in Eq. 1.2, its Laplacian is defined by

$$g_{x^2y^2} = g_{x^2} + g_{y^2}, \tag{1.5}$$

where

$$g_{x^2} = \left(\frac{x^2 - {\sigma_x}^2}{{\sigma_x}^4}\right) \cdot g(x,y;\sigma_x,\sigma_y), \tag{1.6}$$

and

$$g_{y^2} = \left(\frac{y^2 - {\sigma_y}^2}{{\sigma_y}^4}\right) \cdot g(x,y;\sigma_x,\sigma_y). \tag{1.7}$$

1.4.4 Energy minimization

Let $\mathscr{L}$ be the set of binary blobs, and $\mathscr{E}$ be the set of elements in the binary image. Energy minimization finds a labeling f that assigns each element $e \in \mathscr{E}$ to a label $l_e \in \mathscr{L}$, where energy function $\mathbf{E}(f)$ has the minimum.

$$\mathbf{E}(f) = \sum_{e \in \mathscr{E}} D(e, \ell_e) + \sum_{\{e,e'\} \in \mathscr{N}} d(e,e') \cdot \delta(\ell_e \neq \ell_{e'}) \tag{1.8}$$

The term D is the data cost, d is the smoothness cost, and δ is an indicator function. Data cost is the cost of assigning element e to label l_e. $D(e,\ell_e)$ is defined to be the Euclidean distance between the centroid of the element e and the nearest neighbour pixel in blob l_e. Smoothness cost is the cost of assigning neighbouring

elements to different labels. Let $\mathscr{N}$ be the set of nearest element pairs. Then $\forall\{e,e'\} \in \mathscr{N}$,

$$d(e,e') = \exp(-\beta \cdot d_e(e,e')) \tag{1.9}$$

where $d_e(e,e')$ is the Euclidean distance between the centroids of the elements e and e', and β is defined as

$$\beta = (2\left\langle d_e(e,e')\right\rangle)^{-1} \tag{1.10}$$

$\langle\cdot\rangle$ denotes expectation over all pairs of neighbouring elements [12] in an image. $\delta(\ell_e \neq \ell_{e'})$ is equal to 1 if the condition inside the parentheses holds and 0 otherwise.

2 Results and discussion

We evaluate the proposed page segmentation, text line segmentation and word spotting models on various datasets. This section reports the results under a subsection for each of the tasks. Each subsection first provides an abstract description of the employed models, datasets and performance measures. Then, it presents the main results and a discussion on their significance in the related field. Detailed description of the models, datasets and performance measures is given in the articles section.

2.1 Page segmentation

We propose a method [40] to formulate page segmentation task as a dense prediction problem where an FCN predicts the class of each pixel of the input image. The proposed method [40] is evaluated on a dataset that consists of 38 document images from seven different historical Arabic manuscripts. The task is to segment side texts and main texts from non-binary historical manuscripts with complex layout. The segmentation accuracy is evaluated by f-measure metric that shows how accurate the relevant pixels were predicted, while ensuring no pixel is falsely predicted.

Table 2.1 presents the performance of the proposed method [40] compared with another supervised learning method Bukhari et al. [13] which classifies connected components according to hand crafted features and accompanied by a post-processing. Our method [40] performs on par with Bukhari et al. [13] and closes up the gap with results obtained by post-processing (Figure 2.1). We can see that

post processing is quite helpful towards the misclassified side text components due to the class imbalance problem as the number of side text components are less than the number of main text components. More importantly, component classification is not desirable due to redundant and expensive computation of overlapping areas in addition to binary input requirement.

Table 2.1: F-measure comparison for page segmentation between the proposed method [40] and Bukhari et al. [13].

Method	Main Text	Side Text
Bukhari et al. [13]	**0.95**	**0.95**
Proposed method [40]	**0.95**	0.80

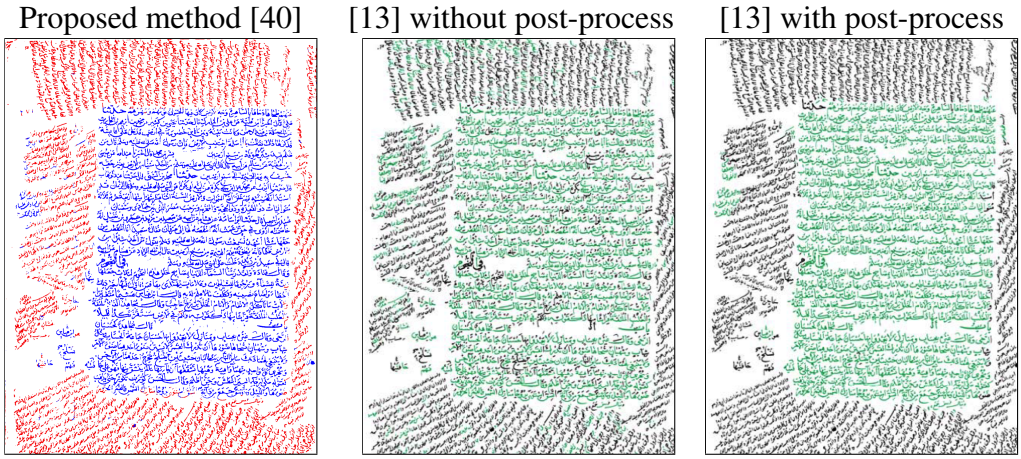

Figure 2.1: Qualitative page segmentation results from the proposed method [40] and Bukhari et al. [13].

2.2 Text line segmentation

We propose a method [38] that formulates text line detection task as a dense prediction of blob lines that strike through the main body area of the characters. The proposed method [38] is evaluated on a dataset that consists of 30 document images from two different historical Arabic manuscripts. The task is to detect the

blob lines that strike through the main and side text lines. The detection accuracy is evaluated by precision that penalizes under-segmentation, recall that penalizes over-segmentation and f-measure that penalizes both.

Table 2.2 presents the performance of the proposed method [38] compared with the performance of a learning free method Cohen et al. [18] which is based on Laplacian of anisotropic Gaussian filters with automatic scale selection. Our method [38] outperforms the method of Cohen et al. [18] while both have lower precision values than recall values. This demonstrates that both methods tend to under-segment. The proposed method [38] under-segments mostly the side text lines due to small number of training patches with side text relative to the number of training patches with main text. Whereas Cohen et al. [18] tend to connect the neighbouring main text blob lines and side text blob lines because it has a post-processing stage for merging the broken blob lines (Figure 2.2).

Table 2.2: Performance comparison for text line detection between the proposed method [38] and Cohen et al. [18].

Method	Recall	Precision	F-measure
Cohen et al. [18]	0.74	0.60	0.66
Proposed method [38]	**0.82**	**0.78**	**0.80**

The previous results show that deep learning relies heavily on the availability of exhaustive data annotations, limiting its performance on side text segmentation. Hence, we introduce a learning free approach that enhances text lines by the Laplacian of multi-oriented anisotropic Gaussian filter. The proposed method [10] is evaluated on VML-MOC (Visual Media Lab-Multiply Oriented and Curved) dataset [10] that consists of 30 document images from multiple manuscripts. The task is to extract heavily skewed and curved side text lines. The extraction accuracy is evaluated by Pixel Intersection over Union (PIU) that shows how accurate the foreground pixels are extracted, and Line Intersection over Union (LIU) that shows how accurate the text lines are detected.

Table 2.3 compares the performance of the proposed method [10] with performance of Cohen et al. [18] that is based on Laplacian of single-oriented anisotropic

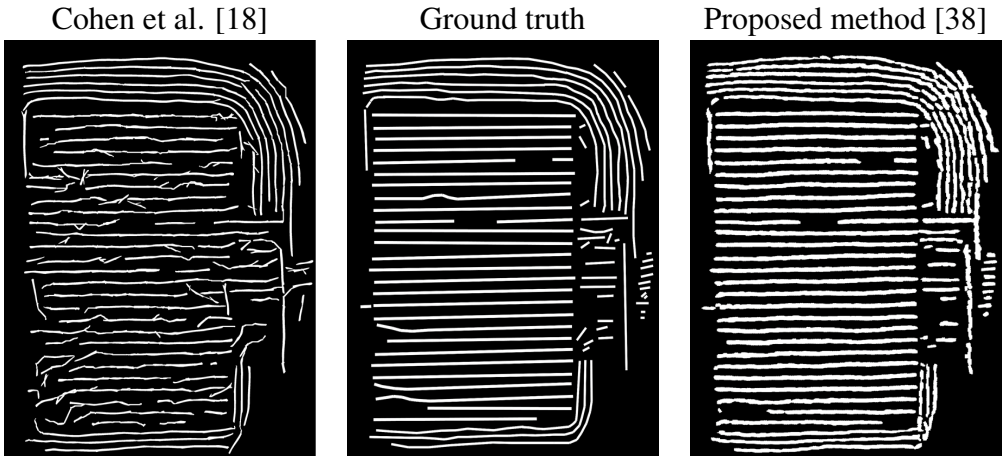

Figure 2.2: A sample ground truth image and corresponding qualitative text line detection results from the proposed method [38] and Cohen et al. [18].

Table 2.3: Performance comparison for text line extraction between the proposed methods and Cohen et al. [18] on VML-MOC dataset.

Method	Line IU	Pixel IU
Cohen et al. [18]	37.88	12.89
Proposed method [10]	**60.99**	**80.96**
Proposed FCN+EM [35]	25.04	48.71
Proposed FCN+EM+Aug [35]	35.12	60.97

Gaussian. The results show that the proposed method [10] surpasses Cohen et al. [18]. This is likely due to multi-oriented Gaussian that can capture the text lines at any orientation, whilst single-oriented Gaussian assumes almost horizontal text lines (Figure 2.3).

Text line detection defines the text lines loosely by baselines or blob lines. On the other hand, text line extraction is a harder task which defines text lines precisely by pixel labels or bounding polygons. We propose a text line extraction method (FCN+EM) which uses Fully Convolutional Network (FCN) to detect blob lines and assists an Energy Minimization (EM) function by these blob lines to extract the text lines (Figure 2.4). The proposed method [35] is evaluated on VML-MOC [10], VML-AHTE (Visual Media Lab-Arabic Handwritten Text line Extraction)

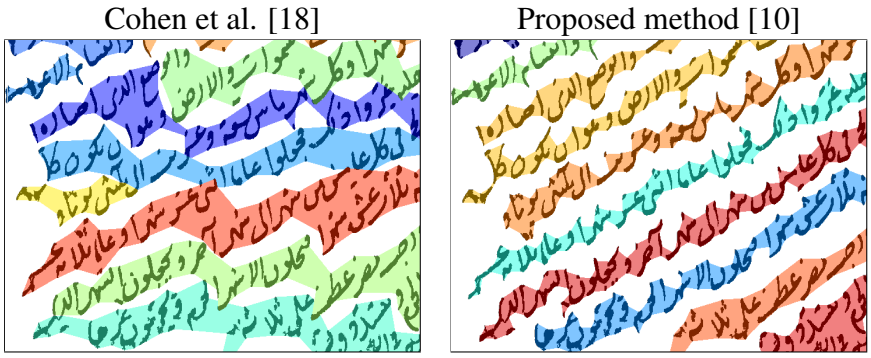

Figure 2.3: Qualitative text line extraction results from the proposed method [10] and Cohen et al. [18].

[35] and DIVA [61] datasets.

Table 2.4 compares the performance of the proposed method [35] with the performance of an instance segmentation method MRCNN [27], on VML-AHTE dataset. Table 2.5 compares the performance of the proposed method [35] with the performance of competition winner [61] on DIVA dataset. We observe that FCN+EM method is ready for direct use without hand crafted features, and it performs best if there are no class imbalance and data scarcity problems. However, neither FCN+EM method nor its augmented version FCN+EM+Aug method are effective on VML-MOC that is a smaller dataset with hard examples (Table 2.3). The probable reason is that data size is not enough for the machine to discriminate among rich variants of the sample instances.

Table 2.4: Performance comparison for text line extraction between proposed methods and a supervised method [35] on VML-AHTE dataset.

Method	LIU	PIU
Unsupervised		
Proposed UTLS [36]	**98.55**	88.95
Proposed method [39]	90.94	83.40
Supervised		
MRCNN [35]	93.08	86.97
Proposed FCN+EM [35]	94.52	**90.01**

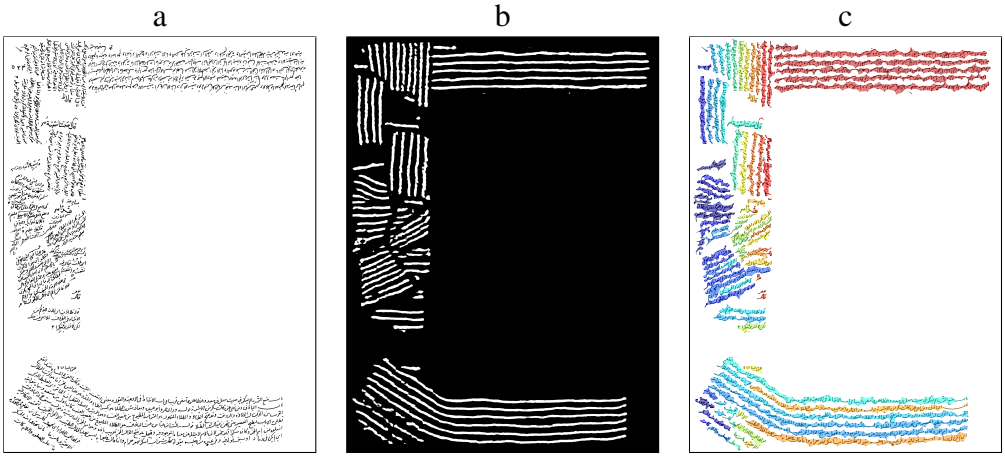

Figure 2.4: Given a handwritten document image (a), FCN detects blob lines that strike through text lines (b). EM with the assistance of detected blob lines extracts the text lines (c).

Learning based methods require a vast amount of annotation effort which consumes time not less than carefully designed features. Hence, we proposed an unsupervised deep learning method for text line segmentation (UTLS) [36]. According to the Gestalt principle [29], proximity and similarity among the elements form the basis of unsupervised segmentation of text lines (Figure 2.5). Inspired by this fact, we train a two branch CNN to learn that two document image patches with relatively same/distinct number of foreground pixels, are similar/different.

Age	5	7	10	11
Perceived text lines				

Figure 2.5: Human visual system can easily perceive handwritten text lines according to proximity and similarity relevance of the elements [29]. Children can segment the text lines although written in a language they are not its readers.

Table 2.5: Performance comparison for text line extraction between the proposed methods and ICDAR2017 competition winner [61] on DIVA dataset.

	CB55		CSG18		CSG863	
Method	LIU	PIU	LIU	PIU	LIU	PIU
Unsupervised						
Winner [61]	98.04	96.67	96.91	96.93	98.62	**97.54**
Proposed UTLS [36]	80.35	77.30	94.30	95.50	90.58	89.40
Proposed method [39]	93.45	90.90	97.25	96.90	92.61	91.50
Supervised						
Proposed FCN+EM [35]	**100**	**97.64**	97.65	97.79	**100**	97.18
Proposed method [33]	**100**	97.04	**98.38**	**98.14**	**100**	97.31

Table 2.4 and Table 2.5 compare the performance of UTLS [36] with the performances of supervised and unsupervised methods on VML-AHTE and DIVA datasets. UTLS [36] results are quite below the supervised methods. The main challenge is the wide spaces between the words in a text line. The wider the space between words, the more likely the algorithm detects it as an inter-line space instead of a text line, which in turn leads to over segmentation. The results get severe when the documents contain text lines with heights that are much greater than or much less than the fixed patch size (Table 2.6). Patch size is a parameter that requires to be hand adjusted for each dataset therefore, UTLS cannot be successful on a heterogeneous dataset that contains document images with various text line heights.

To overcome the aforementioned hand tuning problem of UTLS [36], we propose another unsupervised deep learning algorithm [39] for text line segmentation. The method trains a deep network to answer whether two document image patches contain the same coarse text line pattern or different coarse text line pattern.

Table 2.4, Table 2.5, and Table 2.6 compare the performance of the proposed method [39] with the performances of supervised and unsupervised methods on VML-AHTE, DIVA and ICFHR2010 datasets. The intense touching components among the text lines confuse the statistical supervision of the UTLS method [36]. The proposed method [39] is successful no matter how the touching components

Table 2.6: Performance comparison for text line extraction between the proposed methods, ICFHR2010 competition winner [23], and a supervised method [20] on ICFHR2010 dataset.

Method	DR	RA	FM
Unsupervised			
Winner [23]	**97.54**	**97.72**	**97.63**
Proposed UTLS [36]	73.22	72.38	72.36
Proposed method [39]	71.70	70.46	71.04
Supervised			
Diem et al. [20]	97.18	96.94	97.06

fill the interline space as long as the global pattern of the document is monotonic (Table 2.4, 2.5). While it is true that the proposed method [39] eliminates parameter tuning for each dataset as all experiments have been executed with the same configuration, it is also true that it fails with the sudden changes in the global pattern. Therefore, as one would expect the method is not successful on the documents with fluctuating text lines (Table 2.6).

We tackle the problem of document images with heterogeneous text line characteristics further and propose a method based on automatic scale selection using Laplacian of anisotropic Gaussians [33]. Our findings support that automatic scale selection helps better detect text lines even in the presence of very large variance in text line heights and interline spaces (Table 2.7).

Table 2.7: Performance comparison for text line detection between the proposed method [33], ICDAR2017 competition winner [19] and another recent work [3] on cBAD dataset.

Method	Precision	Recall	F-measure
Learning-free			
Aldavert et al. [3]	74.70	92.60	82.70
Proposed method [33]	82.09	91.13	86.38
Learning-based			
Winner [19]	**97.30**	**97.00**	**97.10**

2.3 Word spotting

Research in historical document analysis is a challenging problem. Benchmark datasets lie at the heart of development, assessment and comparison of the algorithms. We have introduced the Pinkas dataset that is the first dataset in medieval handwritten Hebrew and fully labeled at word, line and page level by a Hebrew paleographer (Figure 2.6). The presented Pinkas dataset will allow to transform research for analyzing Hebrew handwriting. Moreover, such dataset will also be useful to prove the generalizability and robustness of document processing algorithms. We perform baseline experiments with three methods for word spotting, PHOCNet[62], siamese CNN [32], and exemplar SVM[4] to serve as a benchmark for future studies. Results (Table 2.8) show that there is a big room for improvement and confirm the challenging nature of the dataset.

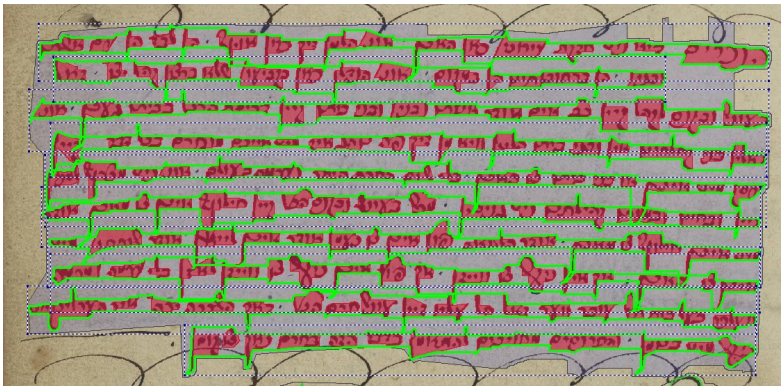

Figure 2.6: Visualization of main text, line and word level segmentation labels. We show main text by purple, lines by green and words by red.

Table 2.8: mAP results of various word spotting methods on the Pinkas dataset.

Dataset	Siamese CNN	PHOCNet	Exemplar SVM
Pinkas	**61.5**	56.6	1.5

3 Conclusion

This thesis has presented several deep learning models to solve page segmentation, text line segmentation and word spotting problems in the field of historical document image analysis. Our findings show that deep learning models can inherently handle the problems arising from complex layout and heterogeneity of documents. However, training deep learning models relies on the availability of exhaustive data annotations which limit their performance in the rare cases where curved lines and fluctuating lines exist, for instance. Learning-free methods outperform deep learning models in such rare cases with scarce training data. On the other hand, unsupervised deep learning methods, which do not require manual data annotation, can be considered successful only on monotonic datasets, as they do comply with the surrogate class labeling. Intuitively, labeling effort is favorable over manually designing feature extractors because the former can be accomplished by human recognition skills, whereas the latter requires further domain expert skills.

Future work would be based on constructing large-scale document image datasets that can offer pre-training opportunity through the full network, instead of originating earlier in the network, as required with the ImageNet dataset, because it is significantly different from document image datasets.

Bibliography

[1] Akiyama, T. and Hagita, N. (1990). Automated entry system for printed documents. *Pattern recognition*, 23(11):1141–1154.

[2] Alberti, M., Vögtlin, L., Pondenkandath, V., Seuret, M., Ingold, R., and Liwicki, M. (2019). Labeling, cutting, grouping: an efficient text line segmentation method for medieval manuscripts. *arXiv preprint arXiv:1906.11894*.

[3] Aldavert, D. and Rusiñol, M. (2018). Manuscript text line detection and segmentation using second-order derivatives. In *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*, pages 293–298. IEEE.

[4] Almazán, J., Gordo, A., Fornés, A., and Valveny, E. (2014). Segmentation-free word spotting with exemplar svms. *Pattern Recognition*, 47(12):3967–3978.

[5] Antonacopoulos, A. (1998). Page segmentation using the description of the background. *Computer Vision and Image Understanding*, 70(3):350–369.

[6] Arivazhagan, M., Srinivasan, H., and Srihari, S. (2007). A statistical approach to handwritten line segmentation. *Document Recognition and Retrieval XIV, Proceedings of SPIE, San Jose, CA*, pages 6500T–1.

[7] Asi, A., Cohen, R., Kedem, K., El-Sana, J., and Dinstein, I. (2014). A coarse-to-fine approach for layout analysis of ancient manuscripts. In *Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on*, pages 140–145. IEEE.

[8] Asi, A., Saabni, R., and El-Sana, J. (2011). Text line segmentation for gray scale historical document images. In *Proceedings of the 2011 workshop on historical document imaging and processing*, pages 120–126.

[9] Bar-Yosef, I., Hagbi, N., Kedem, K., and Dinstein, I. (2009). Line segmentation for degraded handwritten historical documents. In *2009 10th International Conference on Document Analysis and Recognition*, pages 1161–1165. IEEE.

[10] Barakat, B. K., Cohen, R., Rabaev, I., and El-Sana, J. (2019). VML-MOC: Segmenting a multiply oriented and curved handwritten text line dataset. In *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*, volume 6, pages 13–18. IEEE.

[11] Belongie, S., Malik, J., and Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE transactions on pattern analysis and machine intelligence*, 24(4):509–522.

[12] Boykov, Y. Y. and Jolly, M.-P. (2001). Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In *Proceedings eighth IEEE international conference on computer vision. ICCV 2001*, volume 1, pages 105–112. IEEE.

[13] Bukhari, S. S., Breuel, T. M., Asi, A., and El-Sana, J. (2012). Layout analysis for arabic historical document images using machine learning. In *Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on*, pages 639–644. IEEE.

[14] Bukhari, S. S., Shafait, F., and Breuel, T. M. (2009). Script-independent handwritten textlines segmentation using active contours. In *2009 10th International Conference on Document Analysis and Recognition*, pages 446–450. IEEE.

[15] Chen, K. and Seuret, M. (2017). Convolutional neural networks for page segmentation of historical document images. *arXiv preprint arXiv:1704.01474*, pages –.

[16] Chen, K., Seuret, M., Liwicki, M., Hennebert, J., and Ingold, R. (2015). Page segmentation of historical document images with convolutional autoencoders. In *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on*, pages 1011–1015. IEEE.

[17] Chen, K., Wei, H., Hennebert, J., Ingold, R., and Liwicki, M. (2014). Page segmentation for historical handwritten document images using color and texture features. In *Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on*, pages 488–493. IEEE.

[18] Cohen, R., Dinstein, I., El-Sana, J., and Kedem, K. (2014). Using scale-space anisotropic smoothing for text line extraction in historical documents. In *International Conference Image Analysis and Recognition*, pages 349–358. Springer.

[19] Diem, M., Kleber, F., Fiel, S., Grüning, T., and Gatos, B. (2017). cBAD: ICDAR2017 competition on baseline detection. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, volume 1, pages 1355–1360. IEEE.

[20] Diem, M., Kleber, F., and Sablatnig, R. (2013). Text line detection for heterogeneous documents. In *2013 12th International Conference on Document Analysis and Recognition*, pages 743–747. IEEE.

[21] Droby, A., Barakat, B. K., Madi, B., Alaasam, R., and El-Sana, J. (2020). Unsupervised deep learning for handwritten page segmentation. In *2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pages 240–245. IEEE.

[22] Garz, A., Sablatnig, R., and Diem, M. (2011). Layout analysis for historical manuscripts using sift features. In *Document Analysis and Recognition (ICDAR), 2011 International Conference on*, pages 508–512. IEEE.

[23] Gatos, B., Stamatopoulos, N., and Louloudis, G. (2010). ICFHR 2010 handwriting segmentation contest. In *2010 12th International Conference on Frontiers in Handwriting Recognition*, pages 737–742. IEEE.

[24] Grüning, T., Leifert, G., Strauß, T., Michael, J., and Labahn, R. (2019). A two-stage method for text line detection in historical documents. *International Journal on Document Analysis and Recognition (IJDAR)*, 22(3):285–302.

[25] Gruuening, T., Leifert, G., Strauss, T., and Labahn, R. (2017). A robust and binarization-free approach for text line detection in historical documents. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, volume 1, pages 236–241. IEEE.

[26] Ha, J., Haralick, R. M., and Phillips, I. T. (1995). Document page decomposition by the bounding-box project. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 2, pages 1119–1122. IEEE.

[27] He, K., Gkioxari, G., Dollár, P., and Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969.

[28] Kane, S., Lehman, A., and Partridge, E. (2001). Indexing george washington's handwritten manuscripts. *Center for Intelligent Information Retrieval, Computer Science Department, University of Massachusetts, Amherst, MA*, 1003.

[29] Koffka, K. (2013). *Principles of Gestalt psychology*, volume 44. Routledge.

[30] Krishnan, P., Dutta, K., and Jawahar, C. (2016). Deep feature embedding for accurate recognition and retrieval of handwritten text. In *Frontiers in Handwriting Recognition (ICFHR), 2016 15th International Conference on*, pages 289–294. IEEE.

[31] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105.

[32] Kurar Barakat, B., Alasam, R., and El-Sana, J. (2018a). Word spotting using convolutional siamese network. In *2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*, pages 229–234. IEEE.

[33] Kurar Barakat, B., Cohen, R., Droby, A., Rabaev, I., and El-Sana, J. (2020a). Learning-free text line segmentation for historical handwritten documents. *Applied Sciences*, 10(22):8276.

[34] Kurar Barakat, B., Cohen, R., and El-Sana, J. (2019a). Vml-moc: Segmenting a multiply oriented and curved handwritten text line dataset. In *2019 International Conference on Document Analysis and Recognition Workshops (ICDARW)*, volume 6, pages 13–18. IEEE.

[35] Kurar Barakat, B., Droby, A., Alasam, R., Madi, B., Rabaev, I., and El-Sana, J. (2020b). Text line extraction using fully convolutional network and energy minimization. *International Workshop on Pattern Recognition for Cultural Heritage*.

[36] Kurar Barakat, B., Droby, A., Alasam, R., Madi, B., Rabaev, I., Shammes, R., and El-Sana, J. (2020c). Unsupervised deep learning for text line segmentation. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 3651–3656. IEEE.

[37] Kurar Barakat, B., Droby, A., Kassis, M., and El-Sana, J. (2018b). Text line segmentation for challenging handwritten document images using fully convolutional network. In *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pages 374–379. IEEE.

[38] Kurar Barakat, B., Droby, A., Kassis, M., and El-Sana, J. (2018c). Text line segmentation for challenging handwritten document images using fully convolutional network. In *2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pages 374–379. IEEE.

[39] Kurar Barakat, B., Droby, A., Saabni, R., and El-Sana, J. (2021). Unsupervised learning of text line segmentation by differentiating coarse patterns. In *2021 16th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, volume 1, pages 1355–1360. IEEE.

[40] Kurar Barakat, B. and El-Sana, J. (2018). Binarization free layout analysis for arabic historical documents using fully convolutional networks. *International Workshop on Arabic Script Analysis and Recognition*.

[41] Kurar Barakat, B., Rabaev, I., and El-Sana, J. (2019b). The pinkas dataset. *International Conference on Document Analysis and Recognition*.

[42] Lee, J., Hayashi, H., Ohyama, W., and Uchida, S. (2019). Page segmentation using a convolutional neural network with trainable co-occurrence features. In *2019 International Conference on Document Analysis and Recognition (ICDAR)*, pages 1023–1028. IEEE.

[43] Manmatha, R. and Croft, W. (1997). Word spotting: Indexing handwritten archives. *Intelligent Multimedia Information Retrieval Collection*, pages 43–64.

[44] Manmatha, R., Han, C., Riseman, E. M., and Croft, W. B. (1996). Indexing handwriting using word matching. In *Proceedings of the first ACM international conference on Digital libraries*, pages 151–159. ACM.

[45] Manmatha, R. and Srimal, N. (1999). Scale space technique for word segmentation in handwritten documents. In *International conference on scale-space theories in computer vision*, pages 22–33. Springer.

[46] Mechi, O., Mehri, M., Ingold, R., and Amara, N. E. B. (2019). Text line segmentation in historical document images using an adaptive U-Net architecture. In *2019 International Conference on Document Analysis and Recognition (ICDAR)*, pages 369–374. IEEE.

[47] Mehri, M., Gomez-Krämer, P., Héroux, P., Boucher, A., and Mullot, R. (2013). Texture feature evaluation for segmentation of historical document images. In *Proceedings of the 2nd International Workshop on Historical Document Imaging and Processing*, pages 102–109. ACM.

[48] Moysset, B., Kermorvant, C., Wolf, C., and Louradour, J. (2015). Paragraph text segmentation into lines with recurrent neural networks. In *2015 13th International Conference on Document Analysis and Recognition (ICDAR)*, pages 456–460. IEEE.

[49] Oliveira, S. A., Seguin, B., and Kaplan, F. (2018). dhsegment: A generic deep-learning approach for document segmentation. In *2018 16th International*

Conference on Frontiers in Handwriting Recognition (ICFHR), pages 7–12. IEEE.

[50] Ouwayed, N. and Belaïd, A. (2012). A general approach for multi-oriented text line extraction of handwritten documents. *International Journal on Document Analysis and Recognition (IJDAR)*, 15(4):297–314.

[51] Rabaev, I., Biller, O., El-Sana, J., Kedem, K., and Dinstein, I. (2013). Text line detection in corrupted and damaged historical manuscripts. In *2013 12th International Conference on Document Analysis and Recognition*, pages 812–816. IEEE.

[52] Rath, T. M. and Manmatha, R. (2003). Word image matching using dynamic time warping. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages II–II. IEEE.

[53] Rath, T. M. and Manmatha, R. (2007). Word spotting for historical documents. *International Journal on Document Analysis and Recognition*, 9(2):139–152.

[54] Renton, G., Soullard, Y., Chatelain, C., Adam, S., Kermorvant, C., and Paquet, T. (2018). Fully convolutional network with dilated convolutions for handwritten text line segmentation. *International Journal on Document Analysis and Recognition (IJDAR)*, 21(3):177–186.

[55] Rodriguez, J. A. and Perronnin, F. (2008). Local gradient histogram features for word spotting in unconstrained handwritten documents. *Proc. 1st ICFHR*, pages 7–12.

[56] Rothfeder, J. L., Feng, S., and Rath, T. M. (2003). Using corner feature correspondences to rank word images by similarity. In *Computer Vision and Pattern Recognition Workshop, 2003. CVPRW'03. Conference on*, volume 3, pages 30–30. IEEE.

[57] Saabni, R. and El-Sana, J. (2011). Language-independent text lines extraction using seam carving. In *2011 International Conference on Document Analysis and Recognition*, pages 563–568. IEEE.

[58] Scius-Bertrand, A., Voegtlin, L., Alberti, M., Fischer, A., and Bui, M. (2019). Layout analysis and text column segmentation for historical vietnamese steles. In *Proceedings of the 5th International Workshop on Historical Document Imaging and Processing*, pages 84–89.

[59] Scott, G. L. and Longuet-Higgins, H. C. (1991). An algorithm for associating the features of two images. *Proceedings of the Royal Society of London B: Biological Sciences*, 244(1309):21–26.

[60] Sfikas, G., Retsinas, G., and Gatos, B. (2016). Zoning aggregated hypercolumns for keyword spotting. In *Frontiers in Handwriting Recognition (ICFHR), 2016 15th International Conference on*, pages 283–288. IEEE.

[61] Simistira, F., Bouillon, M., Seuret, M., Würsch, M., Alberti, M., Ingold, R., and Liwicki, M. (2017). ICDAR2017 competition on layout analysis for challenging medieval manuscripts. In *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, volume 1, pages 1361–1370. IEEE.

[62] Sudholt, S. and Fink, G. A. (2016). Phocnet: A deep convolutional neural network for word spotting in handwritten documents. In *2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, pages 277–282. IEEE.

[63] Terasawa, K. and Tanaka, Y. (2009). Slit style hog feature for document image word spotting. In *Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on*, pages 116–120. IEEE.

[64] Vo, Q. N., Kim, S. H., Yang, H. J., and Lee, G. S. (2017). Text line segmentation using a fully convolutional network in handwritten document images. *IET Image Processing*, 12(3):438–446.

[65] Wei, H., Seuret, M., Chen, K., Fischer, A., Liwicki, M., and Ingold, R. (2015). Selecting autoencoder features for layout analysis of historical documents. In *Proceedings of the 3rd International Workshop on Historical Document Imaging and Processing*, pages 55–62. ACM.

[66] Wong, K. Y., Casey, R. G., and Wahl, F. M. (1982). Document analysis system. *IBM journal of research and development*, 26(6):647–656.

[67] Xu, Y., He, W., Yin, F., and Liu, C.-L. (2017). Page segmentation for historical handwritten documents using fully convolutional networks. In *Document Analysis and Recognition (ICDAR), 2017 15th International Conference*. IEEE.

תקציר

מחקרים רבים הוקדשו לניתוח תמונת מסמכים היסטוריים, שכן הבנה אוטומטיתשל מסמכים מתמונות מאפשרת לחוקרים להגיע לנתונים הרלוונטיים בקלות. שיטות קונבנציונליות דורשות ממומחים לחלץ בקפידה התכונות מהתמונות. שיטות אלו מוגבלות ביכולתן לעבד תמונות קלט גולמיות. מודלים של למידה עמוקה מורכבים ממספר רבדים שמסוגלים ללמוד ייצוגי נתונים ספציפיים למשימה, עם רמות הפשטה מרובות ישירות מתמונות קלט גולמיות. מודלים אלה הוכיחו הצלחה יוצאת דופן בתחום ראייה ממוחשבת, ומכאן התחילו להשתמש במודלים של למידה עמוקה לניתוח תמונות מסמכים היסטוריים.

תזה זו כוללת אסופת מאמרים המציעים מספר מודלים של למידה עמוקה לגיבוש מגוון פתרונות לניתוח תמונות מסמכים. ראשית, אנו משתמשים ברשת קונבולוציונלית מלאה לסגמנטציה (fully convolutional neural network) של תמונות מסמכים לאזורים הומוגניים. לאחר מכן, אנו מנסחים בעיית זיהוי שורות טקסט כבעיית חיזוי שורות כתם צפופות העוברות דרך שורות (blobs) טקסט. זיהוי שורות טקסט הוא הגדרה רופפת המספקת רק את המיקומים המרחבים של שורות הטקסט. לפיכך, אנו מרחיבים את זיהוי שורות הטקסט לחילוץ שורות טקסט על ידי מזעור פונקציית אנרגיה המונחית על ידי שורות ה עם זאת, אימון מודלים של למידה עמוקה נשענת על זמינות של - blobs מאגרים גדולים של נתונים, דבר המגביל את ביצועיהם במקרים נדירים, כגון, קווים בעלי עקמומיות משתנה והכללה למקרים חדשים. לכן, אנו חוקרים שימוש של אנאיזוטרופיים מרובי (Laplacians of multi-oriented Gaussians) כיוונים לסגמנטציה שורות טקסט בעלות דרגות עקמומיות שונות וכיוונים מרובים. בנוסף, אנו מציגים שני מודלים של למידה עמוקה בלתי מונחית לסגמנטציה שורות טקסט, אשר אינם דורשים (unsupervised deep learning) תיוג ידני. אנו מציעים גם שיטת למידה מונחית שיכולה (supervised learning)

להסתגל להטרוגניות על ידי בחירה אוטומטית בקנה מידה. לבסוף, אנו מציגים את מאגר הנתונים בשם פנקסים שהוא המאגר הנתונים (Pinkas) הראשון בכתב יד עברי מימי הביניים שמתויג במלואו ברמת המילה, השורה והדף על ידי פליאוגרף מומחה. אנו מבצעים ניסויים בסיסיים בשלוש שיטות לאיתור מילים, כדי לשמש קו השוואה התחלתי למחקרים עתידיים במחקר לניתוח כתב היד בשפה העברית.

אני בראט קוראר ברקאת החתום מטה מצהירה בזאת:

חיברתי את חיבורי בעצמי, להוציא עזרת ההדרכה שקיבלתי מאת מנחה.

החומר המדעי הנכלל בעבודה זו הינו פרי מחקרי מתקופת היותי תלמידת מחקר.

בעבודה נכלל חומר מחקרי שהוא פרי שיתוף עם אחרים, למעט עזרה טכנית הנהוגה בעבודה ניסיונית. לפי כך מצורפת בזאת הצהרה על תרומתי ותרומת שותפי למחקר, שאושרה על ידם ומוגשת בהסכמתם.

תאריך: _____ שם: _____ חתימה: _____

העבודה נעשתה בהדרכת

פרופסור ג'יהאד אל־סאנה

המחלקה למדעי המחשב
הפקולטה למדעי הטבע

אוניברסיטת בן־גוריון בנגב
המחלקה למדעי המחשב

Deep Learning for Historical Document Image Analysis

למידה עמוקה לניתוח תמונת מסמכים היסטורית

מחקר לשם מילוי חלקי של הדרישות לקבלת תואר
דוקטור לפילוסופיה

בראט קוראר ברקאת

הוגש לסינאט
אוניברסיטת בן־גוריון בנגב

אישור המנחה
אישור דיקן בית הספר ללימודי מחקר מתקדמים ע"ש קרייטמן

יולי 2021

באר־שבע

אוניברסיטת בן־גוריון בנגב
המחלקה למדעי המחשב

Deep Learning for Historical Document Image Analysis

למידה עמוקה לניתוח תמונת מסמכים היסטורית

מחקר לשם מילוי חלקי של הדרישות לקבלת תואר
דוקטור לפילוסופיה

בראט קוראר ברקאת

הוגש לסינאט
אוניברסיטת בן־גוריון בנגב

יולי 2021

באר־שבע