Document Image Skew Detection and Correction: A Survey

Barun Biswas[1], Dr. Ujjwal Bhattacharya[2], Bidyut B Chaudhuri[3]
[1]*AKCSIT, UNIVERSITY OF CULCUTTA, India*
[2]*CVPR UNIT, ISI-KOLKATA, India*
[3]*Techno India University, India*

***Abstract*—This article presents somewhat detail discussions on document image skew detection methods available in the existing literature. We first categorise the document images into several types and then we discuss different methods available to detect the skew angle of document images belonging to the individual categories. Various methods of document image skew detection include Hough Transform, Principal Components Analysis, Projection Profile based methods, Nearest Neighbor clustering, Connected Component Analysis, Cross Correlation, Radon Transform etc. Discussions on different standard databases of document image samples that have been used by the authors for experimentation purpose are included in the survey. These databases include NIST Federal Register Document Image Database, UW-I, UW-II, UW-III databases, ICDAR DISEC 2013 database, IFN/ENIT database etc. Here, we consider about eight such standard databases. Moreover, we compare different methods with respect to their accuracy and the respective range of skew angles detected by the concerned methods.**

***Index Terms*—Documet Image Processing, Optical Character Recognition (OCR), hough Transform, PCA, Projection Profile Based Methods (PP), Nearest Neighbor Clustering (NN), Connected Component Analysis(CCA)**

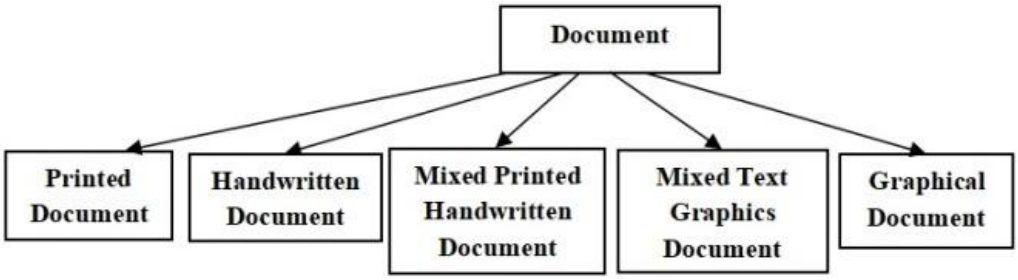
I. INTRODUCTION

Since early 80's a huge number of documents have been stored in computer as soft copy either in machine readable form or as scanned or photographed images. Nowadays digital libraries are established in various organizations. The reasons behind this are many. For example, the volumes of accessible paper documents (hard copies) are often so large that their proper maintenance become difficult or inconvenient. There may arise various situations when hard copy paper documents may get destroyed causing irrecoverable losses. Moreover, there are many valuable but sufficiently old documents physical handling of which may inflict serious damages to them leading to permanent loss of important historical treasures.

In general, an OCR software includes a few preprocessing operations before passing the input document image to the core OCR module. Such operations include binarization [16], skew and slant correction [40], [54][3][66][58] page layout analysis etc. Among these preprocessing operation skew and slant correction have crucial role in the success of the OCR engine and the same is the topic of discussions of this article. Based on the document quality and its architectural complexity, there may be different categories of skewed document image with varying complexities of estimation of the skew. The simplest category of skewed document contains only a single column of text lines. Similarly, a multi-column document may also be skewed in which case skew detection or correction become more difficult. Moreover, there exists some documents containing text lines aligned in different directions and automatic detection of skew (if any) for such documents is more complicated. Other challenges include multilingual documents with scripts of its various languages varying widely or documents containing very little texts while larger area being occupied by drawings, images, charts or some other graphical objects. Also, presence of significant noise in old degraded documents may pose serious problem for detection of skew. The problem may be more serious in cases of handwritten documents. Detection and correction of skew in a document with multiple skews is a great challenge.

The remaining part of this paper is organized as follows. In section II we discuss the problem areas. In section III we discuss about the different methods of

Figure 1: Types of skew in document images



skew and slant correction, in section IV we draw a conclusion about the present methods and finally we have proposed our future work to be done in this field.

II. DATABASE

In this section we discuss about some databases that are publicly available and some database internally generated for the research community. The databases consist of scripts either handwritten or printed. The documents contain figures, diagrams, circuits, tables, architectural plans etc. Some of the reported databases are as follows:

A. The Recognition Improvement Program Set 1 database

In the year 1990 USPS (United States Postal Service) created a dataset The Recognition Improvement Program, Set 1 (RIP-1).

B. The Federal Register data set

A document image database was produced by NIST for evaluating document analysis and recognition technologies and information retrieval systems. This database was known as Special Database 25.

C. The University of Washington English Document Image Data Set I

The UWI data set {R.M. Haralick et. al., UWEnglish document database—(I) manual, Reference Manual,1993} was developed for OCR research and due to its inclusion of real-life noise conditions it has been widely used in the character recognition and document layout detection literature.

D. UW-II English/Japanese Document Image Database

In the series of University of Washington document image databases UW-II is the second one produced by the Intelligent Systems Laboratory.

E. UW-III English/Technical Document Image Database

In the series of document image databases University of Washington UW-III is the third one produced by the Intelligent Systems Laboratory.

F. SRI PAS, HEB, and MAR data sets

About 100 pages scanned at 300 DPI of scripts Pashto, Hebrew, and Marathi were put together by [6]. This is a multi-lingual database.

G. The IFN/ENIT database

For the training and testing of Arabic handwriting recognition software Technical University of Braunschweig has developed The IFN/ENIT-database [88].

H. ICDAR DISEC 2013 database

In 2013 the conference ICDAR organized Document Image Skew Estimation Contest (DISEC 2013). The contest provided a standard database that contains 1550 test image.

I. IAPR TC-11 Dataset

Technical Committee Number 11 (TC11) [105] of International Association for Pattern Recognition (IAPR) takes care of the development of the theory and applications of automatic Reading Systems that recognize or analyse text content in handwritten and printed documents, images, and videos.

III. PROBLEM AREA

At first, we have to categorized the problems so that problems can be solved in a proper way. In this section we try to find different types of problem of skew in document images. We have gone through a huge database and based on the types of the images we classify the document type as:

1. Printed Document

Table 1: List of Database

Year	Author	Name of the Database
	1990 USPS (United States Postal Service)	The Recognition Improvement Program, Set 1 (RIP-1)
1994	NIST: National Institute of Standards and Technology [102]	The Federal Register data set
1993	R.M. Haralick et. al.[90]	The University of Washington English Document Image Data Set I
1993	R.M. Haralick et. al.[90]	UW-II English/Japanese Document Image Database
1996	I.T. Phillips [89]	UW-III English/Technical Document Image Database
2002	Margner et. al.	The IFN/ENIT database
2016	Oussama Zayen [119]	A Dataset for Arabic Text Detection, Tracking and Recognition in News Videos - AcTiV,
2017	Guillaume Chiron [25]	Dataset for the competition on PostOCR Text Correction 2017 (Post-OCR 2017)
2011	Rafael Dueire Lins [63]	Document Image Binarization Platform (DIB Platform),
2009	K. Y. Franke et al [20]	ICDAR 2009 Signature Verification Competition (SigComp2009)
2018	Vincent Christlein [27]	ICDAR 2019 Competition on Image Retrieval for Historical Handwritten Documents Dataset (HisIR19)
2019	Harold Mouch‘ere [73]	ICDAR 2019 Competition on Recognition of Handwritten Mathematical Expressions and Typeset Formula Detec tion (ICDAR2019-CROHME-TDF)
2010	Muhammad Imran Malik [68]	ICFHR 2010 Signature Verification Competition (4NSig-Comp2010)
2019	Yipeng Sun [103]	Arbitrary-Shaped Text (ICDAR-2019 ArT)
2019	Yipeng Sun [104]	Large-scale Street View Text with Partial Labeling (ICDAR-2019 LSVT)
2007	J. Kumar et. al. [57]	TANGO -DocLab web tables from international statistical sites
2014	Harold Mouch‘ere [74]	ICFHR 2014 (CROHME-2014)
2019	Abbas Cheddad [23]	A Swedish Historical Handwritten Digit Dataset (ARDIS)
2014	Joan Andreu S´anchez [106]	Handwritten Text Recognition on tranScriptorium Datasets Bentham R0 (HTR Competition 2014)
2015	Joan Andreu S´anchez [107]	Handwritten Text Recognition on tranScriptorium Datasets: Bentham R1 (HTR Competition 2015) (HTR Competition 2015)
2016	Joan Andreu S´anchez [108]	ICFHR2016 Competition on Handwritten Text Recognition on the READ Dataset (HTR Competition 2016)
2016	Irina Rabaev [92]	Multiply oriented and curved handwritten text line dataset (VML-MOC)
2019	Christian Clausner [28]	RDCL2019 Competition Dataset (Recognition of Documents with Complex Layouts) (RDCL2019

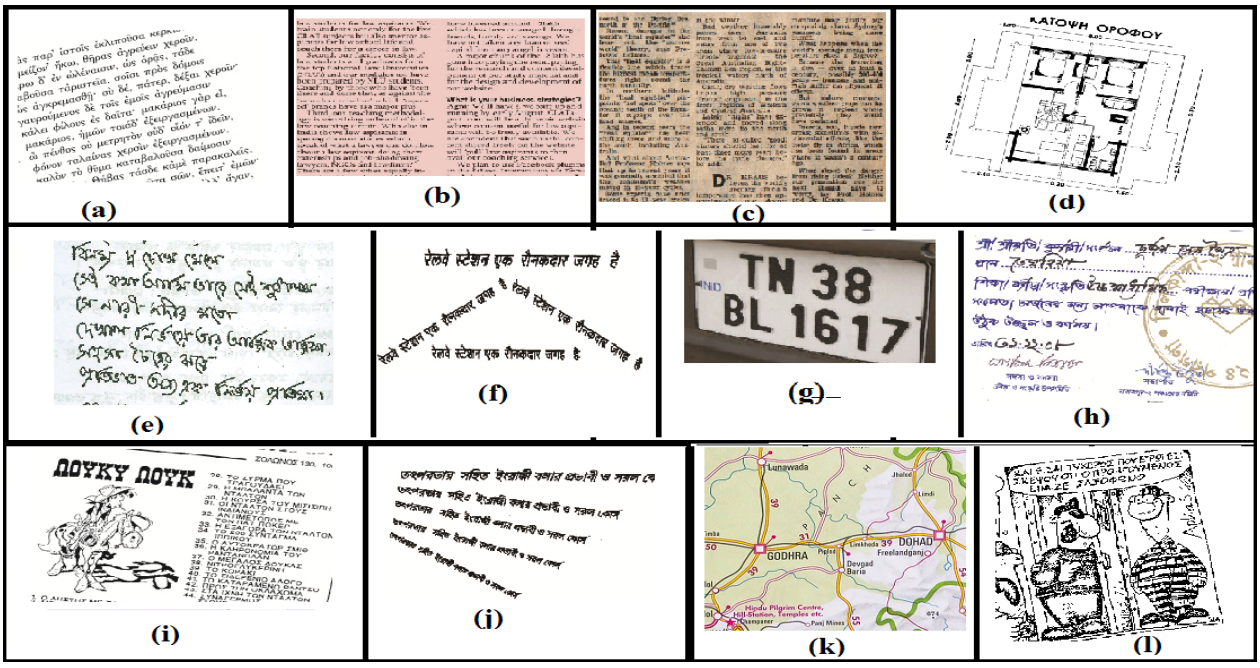


Figure 2: Different types of skewed document image, (a) Single column text skewed document, (b) Double column text skewed document, (c) Multi column text skewed document, (d) Graphical Document, (e) Handwritten manuscript, (f) Multi orientation of text line, (g) Document with perspective distortion (h) Mixed document printed and handwritten document, (i) Mixed text graphics, (j) Document with perspective distortion and multiple orientation, (k) Graphics embedded text of multiple orientation, (l) Text embedded into the graphics

2. Handwritten document
3. Mixed Printed Handwritten Document
4. Mixed Text Graphics Document
5. Graphical Document

The block diagram of figure 3 describe different types of document image. Figure 2 shows examples of different types of skewed image.

A. Printed Document Printed documents are mainly machined printed books, magazine, articles, newspapers, journals etc. In printed documents all text lines are parallel to each other except multi skew artistic documents. When a piece of printed paper is scanned or photographed and saved as image it turns into document image.

B. Mixed Printed Handwritten Document

In these types of documents printed and handwritten scripts are present together. This may be different forms and articles where a previously printed form is filled up by handwritten texts. A document where printed and handwritten both text is present may consider as mixed document. These kinds of documents are seen in different kinds of forms, bank cheque, certificates, answer sheet with printed question, handwritten bills with printed part etc.

C. Mixed Text Graphics Document

These are the document images which we can see outdoor. These images may be any banner, postal, advertisement or any kinds of document images which are captured by camera rather than scanning. There may be perception distortion than skew in the image. The picture can be taken from any angle, for that there may be skew or perception distortion in the image. One example of such problem is vehicle number plate skew correction.

D. Graphical Document

In this section we shall discuss about those papers where skew angle is concerned about the graphical

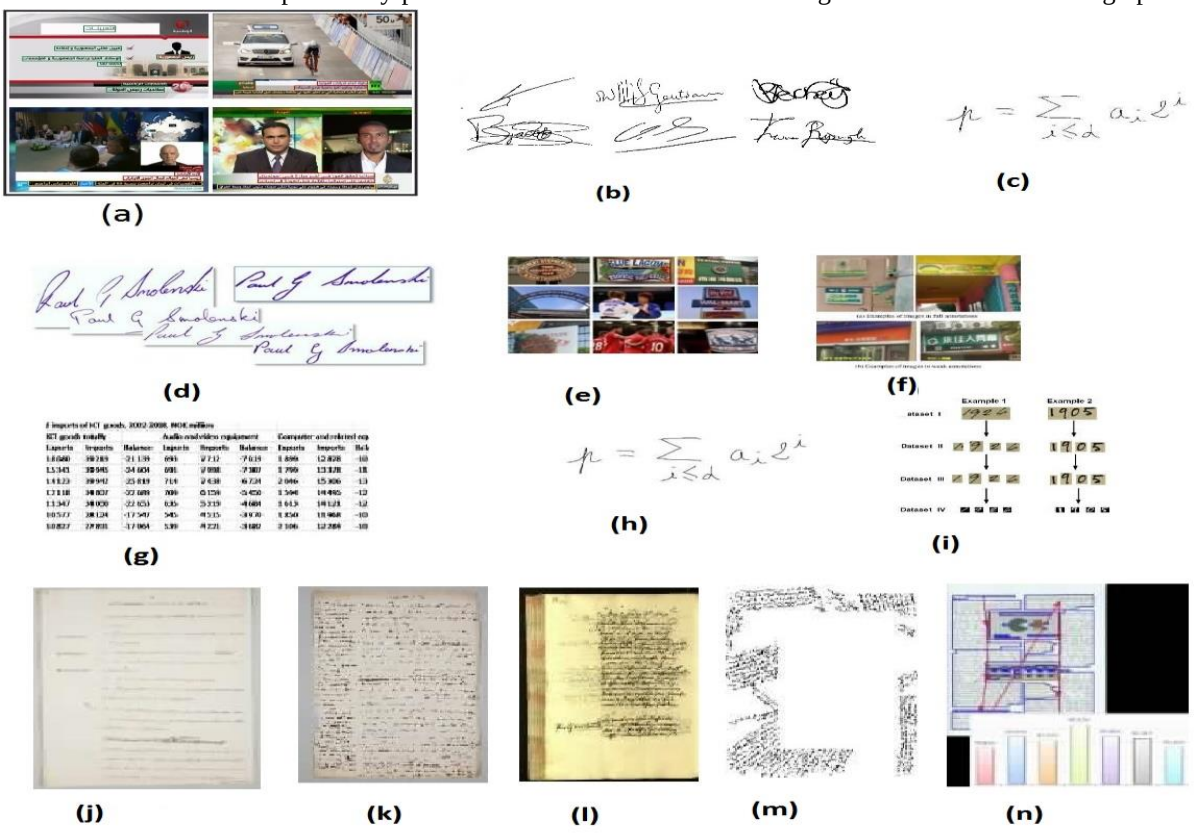


Figure 3: Few example of different types of document in IAPR TC-11 Dataset:a) A Dataset for Arabic Text Detection, Tracking and Recognition in News Videos - AcTiV,[119] b)Dataset for the competition on Post-OCR Text Correction 2017 (Post-OCR 2017) [25] c) Document Image Binarization Platform (DIB Platform) [63] d) ICDAR 2009 Signature Verification Competition (SigComp2009) ,[20] e) ICDAR 2019 Competition on Image Retrieval for Historical Handwritten Documents Dataset (HisIR19) [27] f) ICDAR 2019 Competition on Recognition of Handwritten Mathematical Expressions and Typeset Formula Detection (ICDAR2019-CROHME-TDF) [73] g) ICFHR 2010 Signature Verification Competition (4NSig-Comp2010) [68] h) Arbitrary-Shaped Text (ICDAR-2019 ArT) [103] i) Large-scale Street View Text with Partial Labeling (ICDAR-2019 LSVT) [104] j) TANGO -DocLab web tables from international statistical sites[57] k) ICFHR 2014 CROHME Fourth International Competition on Recognition of Online Handwritten Mathematical Expressions (CROHME2014) [74] l) A Swedish Historical Handwritten Digit Dataset (ARDIS) [23] m) Handwritten Text Recognition on tranScriptorium Datasets Bentham R0 (HTR Competition 2014) [106] n) Handwritten Text Recognition on tranScriptorium Datasets: Bentham R1 (HTR Competition 2015) (HTR Competition 2015) [107] o) ICFHR2016 Competition on Handwritten Text Recognition on the READ Dataset (HTR Competition 2016) [108] p) Multiply oriented and curved handwritten text line dataset (VML-MOC) [92] q) RDCL2019 Competition Dataset (Recognition of Documents with Complex Layouts) (RDCL2019) [28]

image, tabular diagram, engineering drawing or text with images. Skew detection for only text area whether single skew or multi skew is different job than image with text and graphical chart or table or engineering drawing.

IV. AVAILABLE METHODS

In the following paragraph we shall discuss about the different kinds of skew detection and their correction methods of above said problems. We broadly classify the skew detection and correction methods into eight categories. They are:

- Hough Transform
- Principal Components Analysis (PCA)
- Projection Profile Based Methods (PP)
- Nearest Neighbor Clustering (NN)
- Connected Component Analysis(CCA)
- Cross Correlation
- Radon Transform
- Nural Network
- Others

A. Hough Transform

The hough Transform [45] is one of the most popular techniques in image processing to detect a line or a curve in a space [95] [44] [49]. It computes the values for the parameters of all the curves of a particular type (e.g., straight lines) that can pass through each black pixel. Let the distinct point is (x, y) through which the line passes, then the line is represented by:

$$r = x\cos\theta + y\sin\theta$$

Here the θ is the angle that the normal to the line makes with the x axis of the coordinate and r is the distance from origin of the angle to the normal. In practical an accumulator is used to which the value of rand θ is passed as a parameter by a range and depending on their values the decisions are made.

Another approach of skew detection and correction is proposed by Nandini et al. [76] in the year 2008. Two approaches are used, first one is by detecting the centroids of the all text areas and second one is identifying each word as a single blob and finds the orientation of different blobs. After the line representing the skew angle in the document hough Transform is used to detect skew angle. Singh et al. [101] consider BAG (Block Adjacency Graph) to reduce the image pixel before using hough Transform. BAG was first introduced by Yu et al. [116] and used by Yu and Jain [115] to reduce the image pixels before applying the hough Transform to detect skew in the image.

Ishitani [48] , Arulmozhi et al. [15] , Rashid et al. [94] , Pan et al. [83],Le et al. [61], Yin P Y [114] and Aradhya et al.[11], Amin et al. [10] also proposed efficient technique for skew estimation through hough Transform.Kumar et al. [56] presented modified hough Transform for skew estimation.

Malakar et al. [67] uses two-stage hough transform for skew estimation. It is called two stage because in the first stage line are extracted using contour estimation technique [99] and correction of skew using hough transform is applied on the line segment. In the next step each word is extracted using the Spiral Run Length Smearing Algorithm (SRLSA) [100] and skew in the word is corrected (if exist) using hough transform.

B. Principal Components Analysis (PCA)

One of the most using techniques for detecting skew in document images is Principal Components Analysis (PCA). PCA is a way of identifying patterns in data (here in the image), and expressing the data in such a way as to highlight their differences and similarities. It is used in image processing to detect pattern in the image based on which desired features are selected and operation are performed. It is also used for data compression in images. The steps of PCA are as follows:

- Get data of the of the domain area.
- Subtract the mean from the domain area.
- Calculate the covariance matrix
- Calculate the Eigen vectors and eigenvalues of the covariance matrix
- Choosing components and forming a feature vector

The Eigen vector with the heist eigenvalue is the principal components of the data set of the domain area. The operations are performed based on this principal component.

Principal Components Analysis is widely used in detection and correction of vehicle license plate skew [84] [43] [42][24]. The methods used for those purpose are near about same as described above.

C. Projection Profile Based

Projection Profile [46] based process of skew detection [2] is frequently used in image processing.

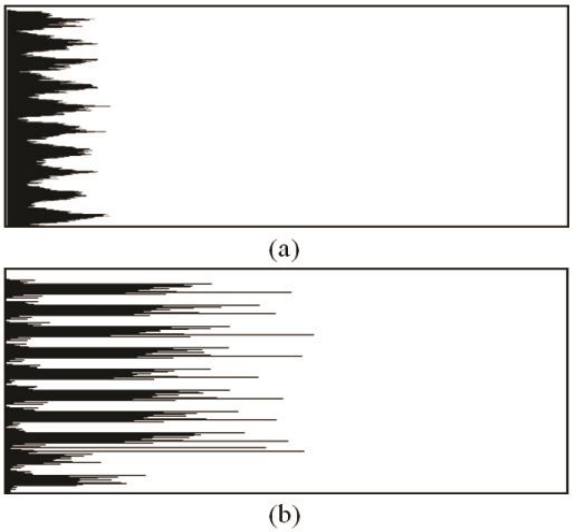


Figure 4: Projection Profile of (a) Skewed document and (b) Non-skewed document

Projection Profile based process uses horizontal and vertical histogram of an image and based on the features of histogram decisions are made about skew angle. In horizontal projection profile analysis each row in one dimensional array stores the number of black pixels; i.e. the foreground pixels, in the original image. The projection profile of an image where skew angle is 0 is a histogram where the peaks of the histograms have the maximum black pixels and valley contain almost 0 numbers of black pixels (as shown in figure 4 (b) and the same of the same skewed image contain more closely peaks and valleys are filled with pixels (as shown in figure 4 (a).

Papandreou et al. [2] proposed a skew detection technique using vertical projection profile approach. The general idea behind using vertical projection is that majority of the Latin alphabets have vertical stroke. So while using vertical projection in the document image a number of peaks at text line position can be noticed but troughs cannot be noticed [34] in case of vertical projection. This feature solves the problem of presence of noise and warp. In this process the image is rotated at an angle and the energy (number of black pixels) at each column is calculated. The higher energy will be matched with the correct skew angle in which the vertical strokes of all rows.

Pan et al. [83] also proposed a paper on skew detection and correction of license plate in the year 2010. To perform the skew estimation, the image is first segmented into five none-overlapping blocks. The local orientation of each block is estimated by gradients of pixels in the block. Local maximum of the direction angle histogram is used to detect the horizontal incline angle of license plate.

D. Nearest Neighbor Clustering

Nearest Neighbor (NN) [47] is an important non-parametric method in the field of pattern recognition. Hashizume et al. [47] in the year 1986 first proposed this method. This method was generalized by O'Gorman [59] in the year1993, in which the Nearest Neighbor clustering was extended to K neighbors for each connected components.

In the year 2001 Liolios et al. [64] apply the K-NN method on the preprocessed document. The proposed algorithm starts from the left-most letter in the top-most of the lines and tries to connect the center of the bounding boxes of all letters in a line in a left to right fashion. K-NN is used on the bounding boxes. In the

year 1999 Jiang et al. [51] K-NN method to calculate the local skew lines and then focused nearest neighbor clustering is applied.

Diem et al. [30] proposed skew estimation technique based on the texts gradient [78] in combination with the classical Focused Nearest Neighbor Clustering [51]. In this method before computing the gradient vector m(x,y) and gradient ve‘ctors angle θ(x, y) smoothing is used with a Gaussian kernel (s = 1.75). For fast computation the authors used Difference of Gaussians (DoG) [38]. Yue et al. [112] also presented an Improved Nearest Neighbor Based Approach to Accurate Document Skew Estimation.

E. Connected Component Analysis

Connected Component labeling [110] is one of the basic tools in computer graphics and computer vision. Labeling connected components is performed by two subsequent raster-scan of the whole binarized or gray scale of color image pixel by pixel in order to identify the connected component regions [5][6]

Connected component labeling for the text area in a document image is used widely in many research papers as a tool to detect skew angle [81] [115] [116] [77] [64] [117] [113] [17]. After detection of connected components in document image many process like hough Transform [81] [115] [116], Nearest Neighbor [64] [117] [113] [17], Projection Profile [65] [21] [1] [51], Principal Component Analysis [77] [84] [43] [42] [24] etc. are used as a final step of skew detection and correction in the document image.

Sarfraz et al. [98] used two algorithm based on histogram statistics and connected components analysis. In the connected component analysis approach within the single line all connected component is selected. The nearest neighbor connected component is dilated and thus each line is thickened. The whole line is assumed as one blob and thus the skew angle is determined. The authors claimed that their proposed algorithm give accuracy more than 99%.

F. Cross Correlation

A gray scale image in the 2D co-ordinate can be represented as following:

$$I(x, y),\ 0 \le x \le X,\ 0 \le y \le Y$$

here X and Y is the extreme value of the co-ordinate x and y. If we take two vertical lines l1($x = x_0$) and l_2($x = x_2$ + d) vertical cross correlation can be represented as:

$$R1(x_0, s) = \sum I(x_0, y)I(x_0 + d, y + s)$$

It is obvious that R_1 is maximized when l_2 is shifted relatively to l_1 at a particular such that the character base lines of two vertical lines are coincide. The correlation of one pair of such lines often produces error prone result, so an accumulated correlation for all pairs which have a distance d as:

$$R(s) = \sum_{x_0}^{x-d} R_1\,(x_0, x)$$

R(s) is calculated by varying s in a predefined region ±S. If S_p is corresponds to the peak in the correlation functions, the skew angle can be determined as:\\

$$\alpha = \arctan \frac{S_p}{d}$$

Here α is the desired skew in the document image.

In the year 1997 Gatos et al. [40] introduce another innovative skew detection and correction process using cross correlation. Before applying the cross correlation (vertical cross correlation in this case) some preprocessing is applied on the document image, they are image smoothing using RLSA [109] and vertical line data acquisition. Latter they detect skew from correlation matrix of two vertical lines and another approach is skew detection from correlation matrix of multiple lines.

G. Radon Transform

Radon transform [80] is function in the 2D special matrix. As the image is represented as 2D matrix in computer or in mathematics we can apply Radon transform on images. The radon transform is a function defined on a straight line, let the line is L, the Rf function is defined as:

$$Rf = \int_L f(x)|\, dx$$

Hence the straight line L with the arc length t can be represented as:

$$(x(t), y(t)) = ((t\sin\alpha + s\cos\alpha), (-t\cos\alpha + s\sin\alpha))$$

Where s is the distance of L from the origin and a is the angle the normal vector to L makes with the x axis. If we use the parameter (s, α) to represent the spaces of line then Radon transform can be represented as:

$$Rf = \int_{-\infty}^{\infty} f(x(t), y(t))|\, yt$$

$$= \int_{-\infty}^{\infty} f\,((ts\sin\alpha + s\cos\alpha), (-t\cos\alpha + s\sin\alpha))dt$$

The variables are defined as:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} s \\ t \end{bmatrix}$$

In the following paragraph we shall discuss about the process of skew estimation and correction using Radon transform. Radon transform takes O(N^2 M) times to execute an document image having size of N×N and skew of M.

Kapoor et al. [53] proposed another method of skew estimation technique using Radon Transform. The proposed algorithm works well both on printed and handwritten document. There are some aspect of handwritten documents, we shall discuss about it in handwritten section. The proposed algorithm works as follows:

1. Radon transform is taken on the document
2. Find the angle corresponding to the highest transform intensity
3. Subtract 90^0 from the detected angle.
4. Shift the angle in anticlockwise direction by Nearest Neighborhood Interpolation Method
5. If the angle is corresponding to the highest intensity 90^0 skew is corrected else go to step iv

The authors claim that the algorithm produces overall accuracy.

In the paper [7] Aithal et al. proposed another method of skew detection and correction method using Radon transform which is much similar as previously discussed methods. In this method they gain the accuracy of 100% within the angle range of $\pm 25^0$Dong et al. [31] used the same method for skew and slant correction. Polar domain information is used in this paper for skew and slant estimation.

In the year 2010 Meng et al. [69] proposed a new novel method of skew estimation using Radon transform. They proposed bagging (bootstrap aggregating) estimator for skew estimation.

H. Nural Network

In the year 2015 Fischer et al. [35] used Convolutional Neural Network for estimating orientation of real image. They considered the problem in three different level. The first level is for the image having orientation of angle $\pm 30^0$, second level is for the image having orientation of angle $\pm 45^0$ and third one is for angle 0^0 - 360^0. To prevent over-fitting they applied image augmentation. For the different level they train different Network namely Net-30, Net-45, and Net-360. For all three tasks they built upon the AlexNet architecture and pretrained on ImageNet.Their architecture consists of 5 convolutional layers, followed by 3 fully connected layers. After each fully connected layer they used rectified linear unit as nonlinearity. At the final stage Normalization and dropout are applied. In all the three Network the result is significantly good and almost obtained 95\% accucary.

Daniel Saez [97] posted on January 12, 2017. The problem here represented as classification problem. He have generate the train data by rotating the image of the database. He have done experiments with two different datasets: the MNIST database of handwritten digits [62] and the Google Street View dataset [118]. NVlabs also uploaded a project of Rotation and skew detection using Deep Learning in GitHub [41]

I. Other Methods

Darko et al. [29] used binary moments for estimating skew angle in the handwritten documents. The moments are sensitive to the rotations [52] and it is applicable to estimate single skew only; i.e. for printed documents [36][37][12]. Alaei et al. [8] used Piece-wise Painting Algorithm (PPA) for skew estimation in the document image. The concept of Painting and Piece-wise Painting Algorithm (PPA) are introduced by Alaei et al. [9]. Chou et al. [26] also used Piece-wise Painting Algorithm (PPA) for skew estimation. They divide the document image into four vertical slabs. The main approach in this process is that the slabs are converted to gray parallelogram for detecting the orientation of the lines and text region. Sharif et al. [33] proposed an algorithm which much similar to Piece-wise Painting algorithm (PPA). Dey et al. [79] also proposed an algorithm based on Piecewise Covering by Parallelogram (PCP) [26]. Robert S. Caprari [96] proposed an efficient algorithm that operates on bit mapped text pattern array to determine the up/down orientation of the page. This algorithm was developed for solving the fax page up/down orientation determination problem

V. DISCUSSION ON PROBLEM SOLVING METHODS

So far we have discussed on different problems including printed and handwritten documents having single [36][37][12] and multi skew [82] [19] [75] [60][32]. Skew angle detection for handwritten documents is not an easy job. For handwritten documents the available algorithms are [39] [55]

[72][91][101][34][4][29][93]. There are some algorithms which are script dependent. Pal et al.[82] proposed the algorithm for Indian script of multi-skew documents, Kapoor et.al. [53] proposed algorithm for Devnagari script, Malakar et al.[67] proposed algorithm for Bangla scripts handwritten documents, Chaudhuri et al.[22] for Indian scripts documents, Jiang et al. [50] and Lu et al. [113] Pan [85] proposed algorithms for Chinese scripts documents, Papandreou [3] for latin scripts. For better results of skew estimation these algorithms are suggested when the document is mainly script dependent. Almost all algorithms work for English documents, so here no algorithms especially for English documents are suggested. There are some other problems where the document images are mainly business card or license plate. These documents are treated as outdoor scene image because these images are taken by the camera. For

Table 3: Comparison of different methods and their range of skew, accuracy and average time

	Skew Angle (Original Image)	Accuracy (%)	Average Time (Seconds)
Projection Profile Technique	1-25 Degrees	99.09	3.02
Principal Component Analysis	1-25 Degrees	98.52	2.98
Peaks and Valleys Analysis	1-25 Degrees	99.09	1.96
Connected Component Analysis	1-25 Degrees	99.5	2.55

these kinds of documents some algorithms are proposed by W. Pan [85], K. Arulmizhi [13] [14][15] J. H. Park [87][86], A. F. Mollah [70][71]. These algorithms are more complex and challenging because the images of license plate and business cards captured by camera sometimes contain background and uneven light and shades. C. Paunwala [67] proposed an algorithm for complex background license plate skew estimation. In most of the cases the images are in gray scale. He Xin-Ping [111] proposed an algorithm which works on colored document images. Noise is another factor which effect the skew detection and correction task more complex. Presence of noise confused the algorithm to choose between the noise and text pixels. Bo et.al. [18] proposed an algorithm which scores a better result for noisy documents. Table 3 shows the Comparison of different methods and their range of skew, accuracy and average time.

REFERENCE

[1] Bagdanov A. and Kanai J. Projection profile based skew estimation algorithm for jbig compressed images. 4th International Conference on Document Analysis. and Recognition (ICDAR), pages 401 – 405, 1997.

[2] Papandreou A. and Gatos B. A novel skew detection technique based on vertical projections. 11th International Conference on Document Analysis and Recognition.(ICDAR), pages 384–388, 2011.

[3] Papandreou. A. and Gatos B. Slant estimation and core-region detection for handwritten latin words. Pattern Recognition Letters, 2012.

[4] Papandreou A. and Gatos B. A coarse to fine skew estimation technique for handwritten words. 12th International Conference on Document Analysis. and Recognition (ICDAR), pages 225–229, 2013.

[5] Rosenfeld A. and Kak A. C. Digital Picture Processing, volume 2. Academic Press, 1982.

[6] Rosenfeld A. and Pfaltz J. Sequential operations in digital picture processing. Journal of the ACM, 13(4):471–494, 1966.

[7] Acharya D. Aithal P. K., Rajesh G. U. and Swamy S. A fast and novel skew estimation approach using radon transform. Int. Journal of Computer Information Systems and Industrial Management Applications, 5:337–344, 2012.

[8] Nagabhushan P. Alaei A., Pal U. and Kimura F. A painting based technique for skew estimation of scanned documents. International Conference on Document Analysis. and Recognition (ICDAR), pages 299–303, 2011.

[9] Pal U. Alaei A. and Nagabhushan P. A new scheme for unconstrained handwritten text-line segmentation. Pattern Recognition, 44(4):917–928, 2011.

[10] A. Amin and S. Fischer. A document skew detection method using the hough transforms. Pattern Analysis and Applications, 3(3):243–253, 2000.

[11] M. V. N. Aradhya, G. Hemantha Kumar, and P. Shivakumara. Skew detection technique for binary document images based on hough transform. World Academy of Science, Engineering and Technology, 32, 2007.

[12] Hematha Kumar G. Aradhya M. and Shivakumara P. Skew estimation technique for binary document images based on thinning and

moments. Journal of Engineering Letters, 14(1):127–134, 2007.

[13] Arumuga P. S. Deepak T. T. Arulmozhi, K. and K. Nallaperumal. A centroids based hough transformation for indian license plate skew detection and correction of ir and color images. IEEE Int. Conf. on Computational Intelligence and Computing Research, pages 1–4, 2012.

[14] Perumal A. S. Mohan V. M. Arulmozhi, K. and K. Nallaperumal. Skew detection and correction of indian vehicle license plate using polar hough transform research. IEEE Int. Conf. on Comp. Intell. and Computing Research, pages 1–4, 2012.

[15] Perumal A. S. Priyadarsini C. S. T. Arulmozhi, K. and K. Nallaperumal. Image refinement using skew angle detection and correction for indian license plates. IEEE Int. Conf. on Computational Intelligence and Computing Research, pages 1–4, 2012.

[16] Gatos B., Pratikakis I., and Perantonis S. J. Adaptive degraded document
image binarization. Pattern Recognition, 39(3):317–327, 2006.

[17] Yuan B. and Tan C. L. Fiducially line based skew estimation. Pattern Recognition, 38:2333 – 2350, 2005.

[18] Yuan B. and Tan C. L. Skew estimation for scanned documents from noises. International Conference on Document Analysis. and Recognition (ICDAR), pages 277–281, 2006.

[19] Kundu M. Nasipuri M. Basu S., Chaudhur C. and Basu D. K. A fast approach to the detection and correction of skew documents. Pattern Recognition, 40(6):1825–1839, 2007.

[20] Franke K. Y. Vuurpijl L. G. Blankers V. L., van den Heuvel C. E. Icdar 2009 signature verification competition.

[21] Sun C. and Si D. Skew and slant correction for document images using gradient direction. 4th International Conference on Document Analysis. and Recognition (ICDAR), pages 142–146, 1997.

[22] B. B. Chaudhuri and U. Pal. Skew angle detection of digitized indian script documents. IEEE Transaction on Pattern Analysis and Machine Intelligence (PAMI), 19(2):182–186, 1997.

[23] Abbas Cheddad. A swedish historical handwritten digit dataset (ardis),(url for download: http://tc11.cvc.uab.es/datasets/ardis1).

[24] Paunwala P. S. Chirag N. and Chaudhary M. An efficient skew detection of license plate images based on wavelet transform and principal component analysis. Int. Conf. on Signal and Image Processing, pages 17 – 22, 2010.

[25] Guillaume Chiron. Dataset for the competition on post-ocr text correction 2017 (post-ocr 2017), (url for download: http://tc11.cvc.uab.es/datasets/post-ocr 20171).

[26] Chu S. Chou C. H. and Chang F. Estimation of skew angles for scanned documents based on piecewise covering by parallelograms. Pattern Recognition, 40:443 – 455, 2007.

[27] Vincent Christlein. Icdar 2019 competition on image retrieval for historical handwritten documents dataset (hisir19), (url for download: http://tc11.cvc.uab.es/datasets/hisir191).

[28] Christian Clausner. Rdcl2019 competition dataset (recognition of documents with complex layouts) (rdcl2019), (url for download: http://tc11.cvc.uab.es/datasets/rdcl20191).

[29] Brodic D. and Milivojevic Z. N. Estimation of the handwritten text skew based on binary moments. Radio engineering, 21(1):162 – 169, 2012.

[30] Kleber F. Diem M. and Sablatnig R. Skew estimation of sparsely inscribed document fragments. 10th IAPR Int. Workshop on Doc. Anal. Sys., pages 292–296, 2012.

[31] Krzy A. Dong J., Dominique P. and Suen C. Y. Cursive word skew/slant corrections based on radon transform. International Conference on Document Analysis. and Recognition (ICDAR), pages 478 – 483, 2005.

[32] Kavallieratou E. and Daskas F. Text line detection and segmentation: uneven skew angles and hill-and-dale writing. Journal of Universal computer Science, 17(1):16–29, 2011.

[33] Sharif A. E. and Movahhedinia N. On skew estimation of persian/arabic printed documents. Journal of Applied Sciences, 8(12):2265 – 2271, 2008.

[34] Pratt F. Secret and Urgent: the Story of Codes and Chiphers. Blue Ribbon Books, 1939.

[35] Dosovitskiy A. Fischer P. and Brox T. Image orientation estimation with convolutional networks. In German Conference on Pattern Recognition (GCPR). Springer, (URL for download: http://lmb.informatik.unifreiburg.de/Publications/2015/FDB15), 2015.

[36] Zitova B. Flusser J. and Suk T. Moments and Moment Invariants in Pattern Recognition. A John Wiley Sons, 2009.

[37] Kapogiannopoulos G. and Kalouptsidis N. A fast high precision algorithm for the estimation of skew angle using moments. In Proceeding of Signal Processing Pattern Recognition and Application SPPRA, pages 275–279, 2002.

[38] Lowe D. G. Distinctive image features from scale invariant key points. Int.Journal of Computer Vision, 60(2):91–110, 2004.

[39] Nicchiotti G. and Scagliola C. Generalized projections: a tool for cursive handwriting normalization. 5th International Conference on Document Analysis. and Recognition (ICDAR), pages 729 – 732, 1999.

[40] Papamarkos N. Gatos B. and Chamzas C. Skew detection and text line position determination in digitized documents. Pattern Recognition, 30(9):1505–1519, 1997.

[41] Github. (url for download: https://github.com/nvlabs/ocrorot).

[42] Shi C. Guo ping W., Min si A. and Hui L. Slant correction of vehicle license plate based on feature point and principal component analysis. Int. Conf. on Computer Science and Software Engineering, pages 487 – 490, 2008.

[43] Shi C. Guo-ping W., Min-si A. and Hui L. Slant correction of vehicle license plate integrates principal component analysis based on color-pair feature pixels and radon transformation. Int. Con. on Computer Science and Software Engineering, pages 919 – 922, 2008.

[44] Ballard D. H. Generalizing the hough transforms to detect arbitrary shapes. Pattern Recognition, 13:111 – 122, 1981.

[45] Ballard D. H. and Brown C. M. Computer Vision. Prentice Hall, New Jersey,USA, 1982.

[46] J. J. Hull. Document image skew detection: Survey and annotated bibliography. Document. Analysis. System. II, World Scientific, pages 40–64, 1998.

[47] Yeh P. S. Hushizumeet A. and Rosenfeld A. A method of detecting the orientation of aligned components. Pattern Recognition Letters, 4(2):125–132, 1986.

[48] Y. Ishitani. Document skew detection based on local region complexity. Second Int. Conf. on Doc. Anal. and Recog.(ICDAR), 10:49 – 52, 1993.

[49] Shen J. and Castan S. An optimal linear operator for step edge detection. CVGIP : Graphical Model and Image Processing, 54(2):121–124, 1992.

[50] Han C. C. Jiang H. F. and Fan K. C. A fast approach to the detection and correction of skew documents. Pattern Recognition Letters, 18:675–686, 1997.

[51] Bunke H Jiang X. and Kljajo D. W. Skew detection of document images by focused nearest-neighbor clustering. 5th International Conference on Document Analysis. and Recognition (ICDAR), pages 629 – 632, 1999.

[52] Hu M. K. Visual pattern recognition by moment invariant. IRE Tran. on Information Theory, 8(2):179–187, 1962.

[53] Bagai D. Kapoor R. and Kamal T. S. A new algorithm for skew detection and correction. Pattern Recognition Letters, 25:1215–1229, 2004.

[54] Fakotakis N. Kavallieratou E. and Kokkinakis G. New algorithms for skewing correction and slant removal on word-level. IEEE 6th Int. Conf. on Electronics, Circuits and Systems, pages 1159–1162, 1999.

[55] Fakotakis N. Kavallieratou E and Kokkinakis G. Skew angle estimation for printed and handwritten documents using the wigner - ville distribution. Image and Vision Computing, 20:813–824, 2002.

[56] D. Kumar and D. Singh. Modified approach of hough transform for skew detection and correction in documented images. International Journal of Research in Computer Science, 2(3):37–40, 2002.

[57] Kasar T. Kumar J. and Ramakrishnan A. G. Edge-based connected component approach for skew correction of complex document images. IEEE TENCON, pages 1–4, 2007.

[58] Jeong S. H. KwagH. K. S., Kim H. and Lee G. S. Efficient skew estimation and correction

algorithm for document images. Image Vision Computation, 20(1):25–35, 2002.

[59] O'Gorman L. The document spectrum for page layout analysis. IEEE Transaction on Pattern Analysis and Machine Intelligence (PAMI), 15(11):1162 – 1173,1993

[60] Spitz A. L. A fast approach to the detection and correction of skew documents. Analysis of compressed document images for dominant skew, multiple skew, and logotype detection, 70(3):321–334, 1998.

[61] D. X. Le and G. R. Thoma. Document skew angle detection algorithm. SPIE Symposium on Aerospace and Remote Sensing - Visual Information Processing II, 1961:251–262, 1993.

[62] Yann LeCun. The mnist database of handwritten digits, (url for download: http://yann.lecun.com/exdb/mnist/.

[63] Rafael Dueire Lins. Document image binarization platform (DIB Platform), 1, ID:DIB Platform 1. http://tc11.cvc.uab.es/datasets/DIB Platform 1. last accessed on 10.05.2021.

[64] Fakotakis N. Liolios N. and Kokkinakis G. Improved document skew detection based on text line connected- component clustering. International Conference on Image Processing, pages 1098 – 1101, 2001.

[65] Fakotakis N. Liolios N. and Kokkinakis G. On the generalization of the form identification and skew detection problem. Pattern Recognition, 35:253 – 264,2002.

[66] Chandran V. Lowther S. and Sridharan S. An accurate method for skew determination in document images. In Proc. of the digital Image Computing Techniques and Applications. Melbourne (Australia), 30(9):1–5, 2002.

[67] S. Malakar, S. Seraogi, R. Sarkar, N. Das, S. Basu, and M. Nasipuri. Twostage skew correction of handwritten bangla document images. 3rd Int. Conf. on Emerging Appl. of Info. Tech. (EAIT), pages 303 – 306, 2012.

[68] Muhammad Imran Malik. Icfhr 2010 signature verification competition (4nsigcomp2010) (4nsigcomp2010), (url for download:http://tc11.cvc.uab.es/datasets/4nsigcomp20101).

[69] Zheng N. Meng G., Pan C. and Sun C. Skew estimation of document images using bagging. IEEE Tran. On Image Processing, 19(7):1837–1846, 2010.

[70] Das N. Sarkar R. Nasipuri M. Mollah A. F., Basu S. and Kundu M. A fast skew correction technique for camera captured business card images. Annual IEEE India Conference (INDICON), pages 1–4, 2009.

[71] Nasipuri M. Mollah A. F., Basu S. and Basu D. K. Text/graphics separation for business card images for mobile devices. in Proc. GREC'09, pages 263–270, 2009.

[72] Yamamoto K Mori S. and Yasuda M. Research on machine recognition of hand printed characters. IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI), 22:386–405, 1984.22

[73] Harold Mouch`ere. Icdar 2019 competition on recognition of handwritten mathematical expressions and typeset formula detection (icdar2019-crohme-tdf), (url for download: http://tc11.cvc.uab.es/datasets/icdar2019-crohme-tdf1).

[74] Harold Mouch`ere. Icfhr 2014 crohme: Fourth international competition on recognition of online handwritten mathematical expressions (crohme-2014), (url for download: http://tc11.cvc.uab.es/datasets/crohme-20142).

[75] Rondel N. Cooperation of multi-layer per-ceptrons for the estimation of skew angle in text document images. International Conference on Document Analysis.and Recognition (ICDAR), pages 1141 – 1144, 1995.

[76] N. Nandini, S. K. Murthy, and G. Hemantha Kumar. Estimation of skew angle in binary document images using hough transform. World Academy of Science, Engineering and Technology, 18:44 – 49, 2008.

[77] Pietikainen M. Okun, O. and J. Sauvola. Robust skew estimation on low resolution document images. Fifth Int. Conf. on Doc. Anal. and Recog.(ICDAR), pages 621 – 624, 1999.

[78] Mahmod R. Omar K., Ramli A. and Sulaiman M. Skew detection and correction of jawi images using gradient direction. Journal of Technology, 37:117–126, 2002.

[79] Dey P. and Noushath S. e-pcp: A robust skew detection method for scanned document images. Pattern Recognition, 43:937 – 948, 2010.

[80] Toft P. The radon transform - theory and implementation. Ph.D. thesis.,Department of Mathematical Modeling, Technical University of Denmark, 1996.

[81] U Pal and B. B. Chaudhuri. An improved document skew angle estimation technique1996 pages. Pattern Recognition Letters, 17:899–904, 1996.

[82] Mitra M. Pal U. and Chaudhuri B.B. Multi-skew detection of indian script documents. International Conference on Document Analysis. and Recognition (ICDAR), pages 292–296, 2001.

[83] R. Pan, X. Ma, and L. Wang. An efficient method for skew correction of license plate. Second Int. Workshop on Education Technology and Computer Science, pages 90 – 93, 2010.

[84] Yan J. Pan, M. and Z. Xiao. An approach to tilt correction of vehicle license plate. int. Con. on Mechatronics and Automation, pages 271 – 275, 2007.

[85] Shi G. Pan W., Jin J. and Wang Q. R. A system for automatic chinese business card recognition. International Conference on Document Analysis and Recognition (ICDAR), pages 577–581, 2001.

[86] Jang I. H. Park J. H. and Kim N. C. Skew correction of business card image inpda. PACRIM, pages 724 – 727, 2003. 23

[87] Jang I. H. Park Z. H. and Kim N. C. Skew correction of business card images acquired in pda. IEEE Proceeding Visual Image Signal Process, 152(6):668 –676, 2005.

[88] M¨argner V. Ellouze N. Pechwitz M., Snoussi M. S. and Amiri H. Ifn/enit—database of handwritten arabic words. Proceedings of the CIFED'02, page 129–136, 2002.

[89] I.T. Phillips. User's reference manual for the uw english/technical document image database iii. uw-iii english/technical document image database manual.1996.

[90] Haralick R. M. Phillips I. T., Ha J and Dori D. The implementation methodology for a cd-rom english document database. International Conference on Document Analysis and Recognition (ICDAR), pages 484–487, 1993.

[91] Plamondon R. and Srihari S. N. On-line and off-line handwritten recognition: a comprehen-sive survey. IEEE Transaction Pattern Analysis and Machine Intelligence. PAMI, 22:62–84, 2000.

[92] Irina Rabaev. Multiply oriented and curved handwritten text line dataset (vml-moc), (url for download: http://tc11.cvc.uab.es/datasets/htr competition 20161).

[93] Piyush M. K. Ramesh D. R. and Mahesh D. D. Skew angle estimation and correction of hand written, textual and large areas of non-textual document images: A novel approach. IPCV, pages 510–515, 2006.

[94] Prati A. Rashid, A. and R. Cucchiara. A real-time embedded solution for skew correction in banknote analysis. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 10:42 – 49,2011.

[95] Duda Richard O. and Peter E. H. Use of the hough transformation to detectlines and curves in pictures. Communications of the ACM, 15:11 – 15, 1972.

[96] Caprari R. S. Algorithm for text page up/down orientation determination.Pattern Recognition Letters, 21:311–317, 2000.

[97] Daniel Saez. Correcting image orientation using convolutional neural networks a hands-on introduction to deep learning applications, (url for download:https://d4nst.github.io/2017/01/12/image-orientation/).

[98] Mahmoud S. A. Sarfraz M. and Rasheed Z. On skew estimation and correction of text. IEEE, Computer Graphics, Imaging and Visualization, 2007.

[99] Halder S. Malakar S. Das N. Basu S. Sarkar, R. and M. Nasipuri. Text line extraction from handwritten document pages based on line contour estimation.Proc. 3rd Int. Conf. on Computing, Communications and Networking Technologies (ICCCNT), pages 26–28, 2012.24

[100] Malakar S. Das N. Basu S. Kundu M. Sarkar, R. and M. Nasipuri. Word extraction and character segmentation from text lines of unconstrained handwritten bangla document images. Journal of Intelligent Systems, 20(3):227–260, 2011.

[101] C. Singh, N. Bhatia, and A. Kaur. Hough transform based fast skew detection and accurate

skew correction methods. Pattern Recognition, 41:3528 – 3546,2008.

[102] Federal stucture. (url for download: https://www.archives.gov/open/datasetfedreg.html).

[103] Yipeng Sun. Arbitrary-shaped text (icdar-2019 art), (url for download:http://tc11.cvc.uab.es/datasets/icdar-2019 art1).

[104] Yipeng Sun. Large-scale street view text with partial labeling (icdar-2019 lsvt), (url for download: http://tc11.cvc.uab.es/datasets/icdar-2019 lsvt1).

[105] TC-11 Reading Systems. TC-11 Online Resources, 2013. [Online; accessed 07/05/2021].

[106] Joan Andreu S´anchez. Handwritten text recognition on transcriptorium datasets: Bentham r0 (htr competition 2014), (url for download:http://tc11.cvc.uab.es/datasets/htr competition 20141).

[107] Joan Andreu S´anchez. Handwritten text recognition on transcriptorium datasets: Bentham r1 (htr competition 2015) (htr competition 2015), (url for download: http://tc11.cvc.uab.es/datasets/htr competition 20151).

[108] Joan Andreu S´anchez. Icfhr2016 competition on handwritten text recognition on the read dataset (htr competition 2016), (url for download: http://tc11.cvc.uab.es/datasets/htr competition 20161.

[109] Casey R. G. Wong K. Y and E. M. Wahl. Document analysis system. IBM Journal Res. Devel, 26(6):647–656, 1982.

[110] Otoo E. Wu K. and Suzuki K. Optimizing two pass connected-component labeling algorithms. Pattern Analysis and Applications, 12(2):117–135, 2009.

[111] Yun-feng L. Xi-ping H. and Qing sheng Z. An efficient algorithm for automatic skew-correction of color document image. Journal of Image and Graphics, 11(3):367–372, 2006.

[112] Lu Y. and Tan C. L. Improved nearest neighbor based approach to accurate document skew estimation. 7th International Conference on Document Analysis. and Recognition (ICDAR), 1:503 – 507, 2003.

[113] Lu Y. and Tan C. L. A nearest neighbor chain based approach to skew estimation in document images. Pattern Recognition Letters, 24:2315–2323, 2003.

[114] P. Y. Yin. Skew detection and block classification of printed documents. Image and Vision Computing, 29(8):567–579, 2001.25

[115] B. Yu and A. K. Jain. A robust and fast skew detection algorithm for generic documents. Pattern Recognition, 29(10):1599–1629, 1996.

[116] B. Yu, X. Lin, Y. Wu, and B. Yuan. Isothetic polygon representation for contours. CVGIP: Image Understanding, 56:264–268, 1992.

[117] Shi Z. and Govindaraju V. Skew detection for complex document images using fuzzy run length. 7th International Conference on Document Analysis. and Recognition (ICDAR), pages 715 – 719, 2003.

[118] Amir Roshan Zamir and Mubarak Shah. Google street view, (url for download: http://crcv.ucf.edu/data/gmcpg eolocalization/).

[119]Oussama Zayene. A dataset for arabic text detection, tracking and recognition in news videos - activ (activ). 2016