Document Blur Detection using Edge Profile Mining

Sajal Maheshwari
IIIT Hyderabad
sajal.maheshwari@students.iiit.ac.in

Pranjal Kumar Rai
IIIT Hyderabad
pranjal.kumarrai@research.iiit.ac.in

Gopal Sharma
IIT Roorkee
gopalsharma1193@gmail.com

Vineet Gandhi
IIIT Hyderabad
vgandhi@iiit.ac.in

ABSTRACT

We present an algorithm for automatic blur detection of document images using a novel approach based on edge intensity profiles. Our main insight is that the edge profiles are a strong indicator of the blur present in the image, with steep profiles implying sharper regions and gradual profiles implying blurred regions. Our approach first retrieves the profiles for each point of intensity transition (each edge point) along the gradient and then uses them to output a quantitative measure indicating the extent of blur in the input image. The real time performance of the proposed approach makes it suitable for most applications. Additionally, our method works for both hand written and digital documents and is agnostic to the font types and sizes, which gives it a major advantage over the currently prevalent learning based approaches. Extensive quantitative and qualitative experiments over two different datasets show that our method outperforms almost all algorithms in current state of the art by a significant margin, especially in cross dataset experiments.

CCS Concepts

•Applied computing → Document management and text processing ;

Keywords

Document image blur detection; Edge analysis; Automated OCR workflows

1. INTRODUCTION

Camera captured document images are becoming more and more prevalent in the digital workflow due to the proliferation of cameras on everyday portable devices like mobile phones. However, there are quality issues with camera captured document images due to reasons like lack of stability during capture process. This turns as a bottleneck in automatic workflows as the non-reliant quality of captured images may lead to the failure of Optical Character Recognition (OCR) algorithms. In this context, automatic assessment of quality of captured document images is extremely useful for numerous applications. For instance, such a tool can reject poor quality images at the input level itself or classify them at the processing stage to limit the manual intervention. A real time quality assessment method like ours can also be suitable for applications to assist the user at the capture time itself [4].

Most of the recent works [22, 12, 8] on automatic quality estimation of document images have resorted to learning based approaches. These approaches formulate the image quality estimation problem as an alternate problem of predicting OCR accuracies, assuming that the input document image quality is correlated with the output OCR accuracies. The usual pipeline is to partition the image into patches and then assign each patch the value of the OCR accuracy obtained over the parent document image and thereafter train a patch level quality predictor. Finally, these approaches perform patch level testing and use the consensus to predict the score of the given test document image. There are two major problems with this pipeline, first the degradations like blur can vary inside a single image itself and hence assigning the same quality factor to each patch in the training image can adversely affect the learning procedure. Second, the patch level analysis may not be consistent with varying font sizes, font types and amount of text present in the patch and hence such methods require testing at multiple scales, which can be computationally expensive.

We address the problem using a learning free approach and propose an Edge Profile Mining (EPM) operator to replace the patch based analysis utilized in most of the previous works. The EPM operator seeks to find the intensity profile for each significant edge point along its gradient (an example is illustrated in Figure 1). We argue that these edge intensity profiles are much stronger and accurate indicators of blur present in the image as compared to a patch based approach, especially in the case of document images where the prominent edges occur only near the text boundaries. As we can observe in Figure 1, a blurred image tends to give elongated edge profiles with feeble gradient magnitudes and a sharper image contains steep profiles with stronger gradient magnitudes.

Edge based analysis for document image assessment has been used in the past as well [21], but most of these methods rely on edge detectors which themselves can fail in the presence of blur. Moreover, even if the edge is detected the analysis is done by selecting patches centred around the edge

ICVGIP, December 18-22, 2016, Guwahati, India

DOI: http://dx.doi.org/10.1145/3009977.3009982

pixels, which may lead to problems associated with the patch based methods. This further highlights the importance of proposed EPM operator, which provides a much more concrete representation of the edge neighbourhood. Another major advantage of edge profile features is that they generalize for varying font sizes (even within the same document) as they can easily be normalized and compared.

Contributions:.

1. We introduce a novel EPM operator, which minimally and accurately captures the important properties such as extent of blur of the region around the character boundaries.

2. We propose a novel framework to exploit the edge profiles obtained using the EPM operator for the task of predicting the blurriness in an image, which we show highly correlates with overall quality and OCR accuracy of the image.

3. We propose a novel Document Blur (DB) dataset containing large variations in font sizes, font types, capture angle, document layout etc. for a more detailed analysis of Document Image Quality Assessment (DIQA) algorithms.

4. We perform extensive qualitative and quantitative experiments to demonstrate the usefulness of the proposed framework, and show significant improvements over the current state of the art techniques.

2. RELATED WORK

Most of the early methods for Image Quality Assessment (IQA) have been designed for natural images and they have been analyzed in [6]. Based on this analysis, an improved IQA algorithm using the concept of Just Noticeable Blur was proposed in [6] and it was later extended in [14] by using cumulative probability of blur detection. Another interesting line of work [12, 13] used the variations in Natural Scene Statistics (NSS) to quantify the quality of an image. Many other low level features based on gradient distributions or spectral/phase response etc. have been used to predict the sharpness or other quality aspects in an image [7, 18]. However, many of these methods may not apply for Document Image Quality Assessment(DIQA), as document image properties are quite different than those of natural images [21]. Additionally, the quality of document images are correlated with the OCR accuracies which provides a reasonable measure for quantitative experiments (in contrast IQA highly relies on human perception).

Due to the above stated reasons, many recent approaches have been proposed specifically for quality assessment of document images. The work by Blando et al. [3] was one of the earliest methods which measured the quality of the image by predicting OCR accuracies. This method relied on low level features like amount of white speckle or character fragments etc. A gradient searching based approach was proposed in [15] for a similar task. More recent methods which rely on low level features include work by Kumar et al. [9] and Rusinol et al. [17]. The work in [9] used the ratio of sharper edge pixels over total number of edge pixels in an image to predict its blurriness. However, their method does not generalize if the image is rescaled. The work in [17] uses a combination of multiple measures based on features like gradient energy, squared energy, histogram range etc. and then utilizes the worst performing measure to predict the overall quality. The measures are computed on patch level and hence their method is susceptible to error with varying font sizes and varying amount on text in the image.

Learning based approaches [20, 22, 23] have shown superior results for predicting image quality compared to those relying on hand crafted low level features. The work in [22] uses raw-image-patches from a set of unlabelled images to learn a dictionary in an unsupervised manner. Later this is used in a regression framework to predict the quality of an image. Similar idea was extended for quality estimation of historical document images in [20]. The major disadvantage of work in [20, 22] was the computational load, which was reduced in [23] using a supervised filter learning approach.

More recently, deep learning based approaches have been extremely successful for the task of feature learning. It has been successful in many core computer vision applications like classification, object detection etc [11]. Recently Kang et al. [8] introduced a deep learning based approach for DIQA. The network is trained on patch level, which when given a patch, outputs its quality score and the mean of multiple patches on the test image is used to predict the quality of the entire image. However, we observed that such methods do not generalize very well over different kind of document images (with varying styles, font types, font sizes etc) due to the over-fitting tendency of the network.

Despite the success of learning based approaches in the recent past, we observed that they are not exploiting the specific knowledge of being working on document image and most of them have been derived from similar work in IQA. In this paper, we propose a novel EPM operator which takes advantage of one simple observation that edges at the character level is the most important cue to identify degradation in an document image and we show that a minimal framework based on the proposed operator outperforms almost all of the existing approaches.

3. BLUR DETECTION ALGORITHM

In this section we explain the proposed blur detection framework. We first introduce the notion of Edge Profile and then describe the proposed Edge Profile Mining (EPM) operator. We then explain the process to select the unimodal edge profiles which are typically useful to judge the image quality (the rest of the edge profiles are ignored during decision process). We then describe the process to judge the quality of the input image, given all the filtered edge profiles.

3.1 EPM Operator

An edge profile is defined as the set of intensity values encountered while moving along the gradient of an edge, starting and ending at near zero gradient magnitudes (the zero crossing occurs exactly in the middle of an edge profile). An example of edge profile is shown in the second column of Figure 1. We argue in this paper that edge profiles are a strong indicator of the blurriness in a document image, as most of the edges occur near the character boundaries, where the extent of blur is often revealed. The edge profiles also facilitate a much more robust framework for blur detection as compared to other edge based techniques relying on off the shelf edge detectors, which may themselves fail in presence

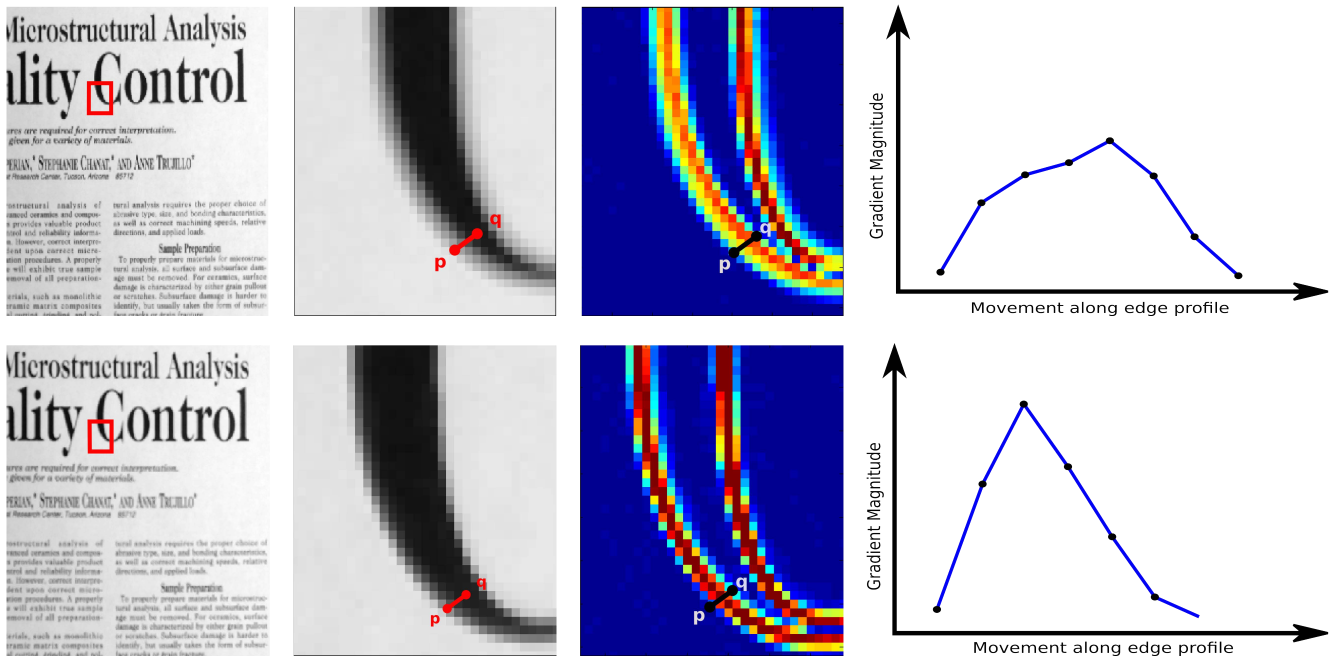


Figure 1: Illustration of an edge profile from a blurred image(top) and a sharp image(bottom). First column shows the original image and the considered area. A zoomed in view of the considered area is shown in second column. The detected edge profile pq starting at position p and terminating at position q is illustrated using the red line. Column 3 illustrates the gradient magnitude values (from blue to red in ascending order). The last column shows the edge profile as moving along the direction of the gradient. We can observe that the edge profile for a sharp image tends to be shorter and steeper as compared to its blurred counterpart.

of significant blur.

The EPM operator seeks to find all edge profiles present in an image. The first step is to compute the gradient magnitude and orientation for each pixel in the input image. The set of pixels $\mathbb{S}$, having a gradient magnitude above a threshold is then considered by the EPM operator. For each given pixel $p \in \mathbb{S}$, the operator moves along the ray $r = p + n * I_d(p)$, where $I_d(p)$ is the gradient orientation at the position p and $n \geq 0$. The operator terminates as it encounters another pixel which does not belong to $\mathbb{S}$ (an example is illustrated in Figure 1 where the process starts at pixel p and terminates at pixel q). All the pixels encountered along the ray are marked as visited and they constitute an edge profile (for example segment pq in Figure 1 is an edge profile). The above process is repeated until there is no unvisited pixel left in $\mathbb{S}$. Finally, the output of EPM operator is a set of all edge profiles $\mathbb{E}_i \in \xi$. The entire process is described in Algorithm 1.

An analogous process has been used to find the width of strokes in [5] for the text detection application in natural images. A major difference is that stroke width transform assigns a label to each pixel indicating the width of the stroke it belongs to whereas we retrieve and store the entire edge profile consisting of gradient magnitudes (apart from the the principal difference between a stroke and an edge profile itself). Interestingly the variation proposed in our work reveals a totally different aspect of the image which as we show, can be extremely useful for many complementary applications like image blur estimation considered in the presented work.

3.2 Unimodal filtering

The edge profiles mined by EPM operator may not always terminate between the stroke of the character as is desired ideally. For instance, in case of extremely narrow characters the edge profile may well encompass the entire stroke. Nevertheless, such a variation possesses similar characteristics as of the ideal edge profile and does not affect the decision process for DIQA. However in cases of extreme blur (where the gradient may not go below the threshold, even inside longer strokes) or closely spaced characters, profiles over multiple strokes may combine during a single pass of the operator (a single edge profile may spread across multiple zero crossings). This leads to what we call a multimodal profile (as illustrated in Figure 3.2) and might be misguiding in the decision process of DIQA. The presence of arbitrary noise and speckle may also lead to edge profiles showing abnormal behaviour which can again lead to erroneous judgement.

We eliminate these unwanted edge profiles using a minimal approach of unimodal Gaussian filtering. This is based on the observation that an ideal unimodal edge profile has its maximum gradient magnitude at the zero crossing, which gradually diminishes symmetrically as one moves along either end of the edge profile. On the other hand, multimodal or other abnormal profile may not exhibit this characteristic. Hence, once a detected profile is transformed to a normalized length it should correlate well with a Gaussian like profile with mean value situated near the middle of the edge profile (ideal zero crossing).

Therefore, we first transform each profile to a normalized length l and correlate it with a Gaussian distribution centred around $l/2$ and standard deviation of $l/2$. If the correlation

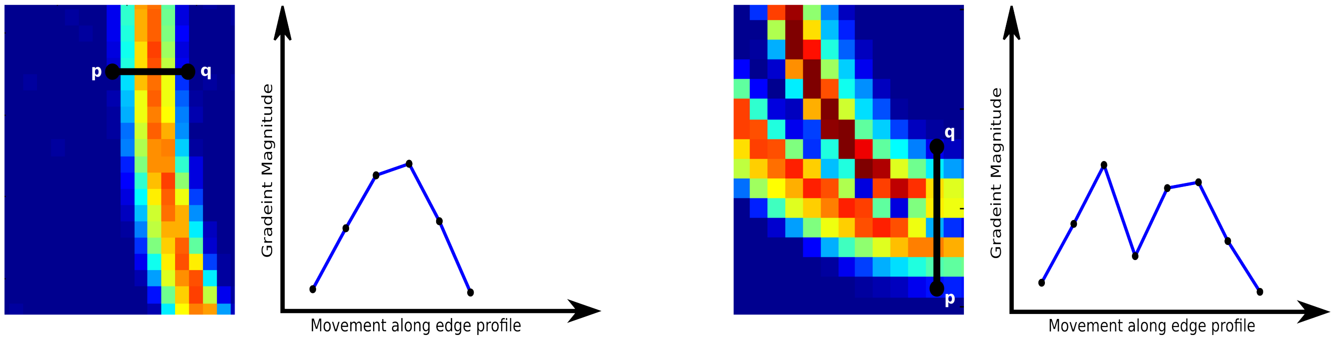


Figure 2: Unimodal vs Multimodal edge profiles. The left figure illustrates the gradient image and the respective unimodal edge profile pq. The right figure illustrates the gradient image and the respective multimodal edge profile pq. Such multimodal profiles may occur in presence of extreme blur.

Algorithm 1 The algorithm of EPM operator

Require: Input Image (I),
  Gradient Magnitude and Oriention Images (I_M, I_D)
  Gradient threshold τ.

  Compute the set $\mathbb{S}$ of pixel positions with $I_M > \tau$.
  Initialize empty set of edge profiles ξ.
  repeat
    Draw a candidate p from $\mathbb{S}$.
    Initialize an edge profile $\mathbb{E}$ with no entries.
    Next candidate pixel $r = p$.
    while true **do**
      Add $I_M(r)$ to $\mathbb{E}$.
      next candidate pixel $r \leftarrow r + nI_d(p), n > 0$
      if $I_M(r) \leq \tau$ **then**
        break
      if $I_M(r) \epsilon \mathbb{S}$ **then**
        Draw r from $\mathbb{S}$
    Add the edge profile $\mathbb{E}$ to ξ.
  until $\mathbb{S}$ is empty
  return set of all edge profiles ξ.

exceeds a threshold T, the edge profile is considered to be appropriate for quality assessment task, else the profile is rejected.

3.3 Calculation of Quality Score

The edge profiles after filtering are mostly unimodal profiles and these strokes/profiles are finally used for predicting the quality of the image. We estimate a quantitative score for each profile and the quality of the parent image is then calculated using a weighted mean of the individual profiles depending on their widths.

We quantify the quality/blurriness of an edge profile using the notion of standard deviation. Such a minimal measure is directly used to quantify the sharpness due to two major observations. First, the minimum value in the sequence of gradient magnitudes in a given edge profile $\mathbb{E}$ is bounded by the threshold τ (occuring either to the start or the end of an edge profile segment). Second, only the prefiltered edge profiles are considered for the decision process, the mean value of the gradient magnitudes also occurs in a reasonable range. Hence, once an edge profile is transformed to a normalized length, the standard deviation is also bounded in a reasonable range and its value directly indicates the blurriness of the corresponding edge profile (higher standard deviation referring to steep and sharper profiles).

The individual scores of the profiles are then used to predict the quality of the entire image. We perform a weighted mean over all filtered edge profiles ξ_f, where the weights correspond to their relative length. The weighted mean shows improvement over the un-weighted case, possibly because it is also encompassing the aspect of length variations in edge profiles (blurred edge profiles tends to be longer than the sharp ones).

4. EXPERIMENTAL RESULTS

In this section we present the experimental results of the proposed framework and compare them with the state-of-the-art learning based as well as metric based approaches assuming the OCR accuracy of a document image as the ground truth representing its quality. We also present a qualitative evaluation of our approach on printed as well as handwritten documents.

4.1 Datasets

(1) **Sharpness-OCR-Correlation(SOC) Dataset** [10] - The SOC Dataset is a publicly available dataset composed of a set of 25 documents. Each document has been acquired 6-8 times at varying focal lengths. The final dataset has a total of 175 images. Each of these images have been OCRed by three softwares(ABBYYFineReader [1], Tesseract [19] and OmniPage [2]). The comparison between OCR outputs and transcription in terms of percentage of character accuracy is used for quantitative evaluation of image quality, which is computed using ISRI OCR [16].

(2) **Document Blur Dataset (DB)** - The SOC dataset has images taken from approximately equal angles and it is also limited in terms of layout variations. In this work, we present a more realistic and challenging dataset with images being captured at different angles to the camera, mimicking a realistic user behaviour. The dataset is composed of 15 documents with varying amount of English machine-printed text across a range of font sizes recognized by OCR engines. We captured 5-8 images of each document altering the focal lengths in a fashion similar to the SOC dataset. The final dataset is composed of 95 colour images of resolution 1024×920.

We then apply standard OCR engines as used in SOC to

	LCC	SROCC
ΔDOM	0.56	0.62
Focus Measure	0.65	0.84
CORNIA	0.88	0.85
DCNN	0.89	0.88
Proposed Approach	**0.74**	**0.78**

Table 1: Comparison of different approaches on SOC dataset.

	LCC	SROCC
ΔDOM	0.21	0.28
Focus Measure	0.28	0.35
CORNIA	0.67	0.60
DCNN	0.88	0.86
Proposed Approach	**0.71**	**0.72**

Table 2: Comparison of different approaches on DB dataset.

recognize the text in an image and then use ISRI evaluation tool to obtain the OCR accuracies for each image (using the ground truth text and the *OCRed* output). We then assign the obtained OCR accuracies to each corresponding image as ground truth for quality evaluation.

4.2 Quantitative Evaluation

The quantitative evaluation of our approach is based on the computation of correlation between the predicted quality scores and ground truth OCR accuracies as done traditionally for both natural and document image quality assessment. Specifically, Linear Cross Correlation(LCC) and Spearman Rank Order Cross Correlaition(SROCC) have been used for evaluation of the proposed algorithm and its comparison with previous approaches. LCC measures the degree of linear dependency between two variables while SROCC is the correlation of rank values of the two variables, assessing the monotonic relationships between them.

We compare the proposed method with four different approaches from the state of the art, namely the ΔDOM [9], Focus Measure (FM) [17], CORNIA [22] and Document Convolution Neural Network (DCNN) [8]. Two of these approaches (ΔDOM and FM) rely on hand crafted low level features and the other two are based on automated filter learning. We first make comparisons with these methods on individual datasets and then present cross dataset results for learning based methods.

The results on SOC dataset for each of the above approaches are listed in Table 1. Here, ΔDOM, Focus Measure and our method were tested on entire dataset while the CORNIA and DCNN were only tested on 20% of the dataset. In DCNN, 60% of the images were used for training and 20% images were used for validation and in CORNIA 60% data was used for training itself. These ratios are taken as per the specifications mentioned in the respective papers. We used the available CORNIA code for the experiments and wrote our own code for DCNN (as the original codes were not available) and obtained almost similar accuracies as claimed in the original papers. The proposed approach outperforms the ΔDOM and FM by at least 10-15%. On the other hand, the learning based approaches outperform our method by a similar margin. However, it should be observed that the testing in these cases is done on a limited part of the dataset only.

	LCC	SROCC
CORNIA	0.43	0.54
DCNN	0.43	0.70
Proposed Approach	**0.71**	**0.72**

Table 3: Comparison with learning based approaches, with training on SOC dataset and testing on DB dataset.

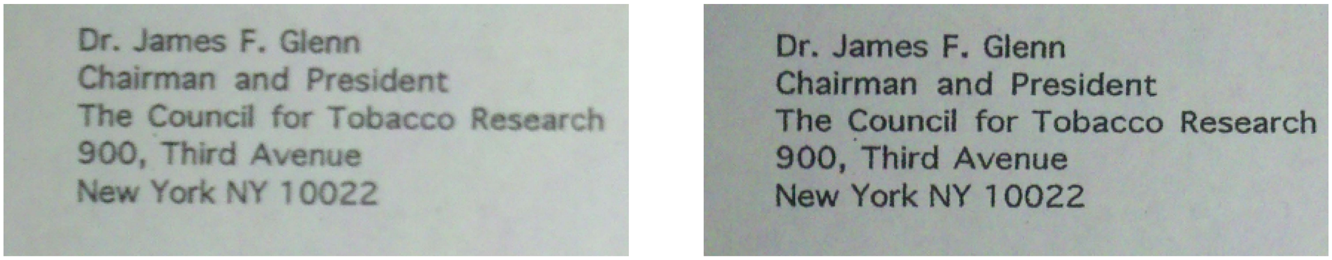
	LCC	SROCC
CORNIA	0.45	0.51
DCNN	0.59	0.74
Proposed Approach	**0.74**	**0.78**

Table 4: Comparison with learning based approaches, with training on DB dataset and testing on SOC dataset.

Similarly, we compare different approaches on the proposed DB dataset and the results are presented in Table 2. We observe that accuracies of almost all methods reduce on the DB dataset, which may be because of the larger variations (in font size, font types, capture angle, layouts etc.) present in this dataset. The observed comparison trends are otherwise similar to the previous case, except that our method outperforms CORNIA over this dataset. The DCNN still remains the best performing approach. Based on the results on the two datasets, this can be established that our method clearly outperforms other approaches based on low level features, but further analysis is needed for comparison with learning based approaches.

More interesting insights start to appear, as we perform cross dataset experiments for learning based approaches i.e. training on one of the datasets and testing on another one. The obtained results with training on SOC dataset and testing on DB dataset are presented in Table 3 and the other case is presented in Table 4. Here the pre-trained models used in previous comparisons were directly used for testing on other dataset. One major difference is that the testing is now performed on the entire dataset. Interestingly, our method now outperforms the DCNN method by almost 30% over LCC in first case and by almost 15% in the second case.

Couple of interesting inferences may be drawn from the set of experiments performed above. First, the proposed approach seems to generalize quite well, which can be inferred by consistency of results over the different datasets, seemingly capturing the desired aspects of blurriness. Second, given the main reason of degradation in both SOC and DB dataset is primarily because of blur, the performance degradation of learning based approaches over cross dataset experiments suggests some form of over fitting, which may be happening when training on 80% and testing on only 20% percent of the data. Another reason for the failure to capture the desired aspects could be the underlying patch based training approach, where the obtained OCR accuracy over the entire document is assigned to all the extracted patches (ignoring the blur variation within a single image). This suggests a need for a better training pipeline to fully exploit



OCR Accuracy : 97.15 %   OCR Accuracy : 98.15 %

Figure 3: Two images with different levels of blur showing similar OCR accuracies.

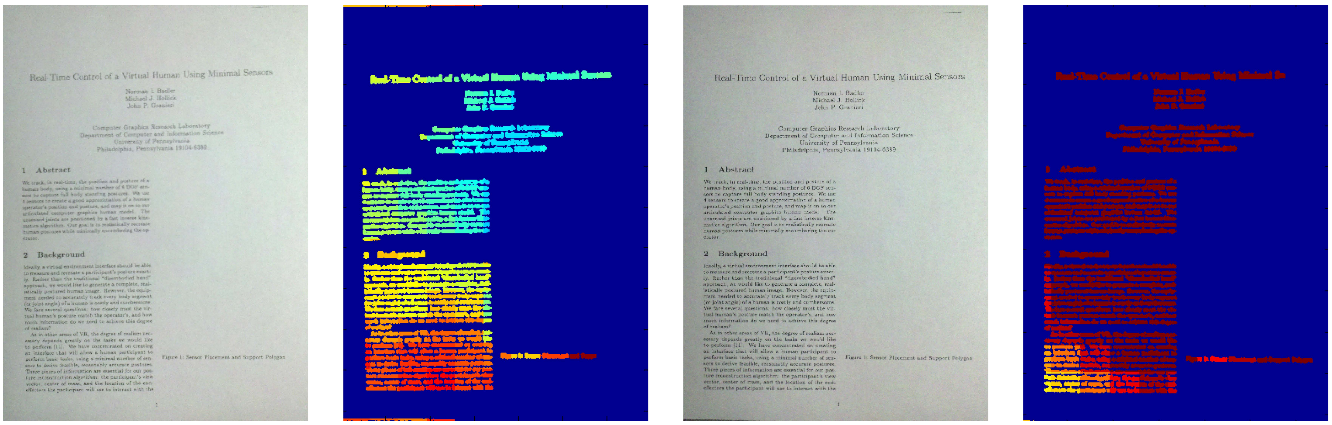


Figure 4: Obtained heat maps using our approach over two different images from SOC dataset. (Red represents the sharp regions and blue represents blurred regions). We can clearly observe the varying amount of blur in both the images.

the true potential of CNN based learning approaches.

4.3 Qualitative Evaluation

In this section, we evaluate our results qualitatively. The first important qualitative observation we made was that the OCR accuracies obtained over a document image do not necessarily reflect the amount of blur or other degradations present in an image. An example is illustrated in Figure 4.2, where we can observe that the image on the left is much more blurred than on the right, however the OCR accuracies are quite high for both. This may actually be one of factors reducing the LCC and SROCC values for our method.

We performed a pilot study to further verify the above observations. We randomly sampled pairs of images from the two dataset and asked 10 different human observers to identify the sharper image between the two. Each observer was shown a set of 50 random pairs and their response was compared with the decision based on the predicted quality score of the proposed algorithm. We found an average agreement of about 93.6% over the 500 decisions made by the observers and our algorithm. This further demonstrates the effectiveness of our approach.

Another advantage of the proposed algorithm is that it can be used for local image quality assessment as well. We demonstrate this by generating pixel level heat maps visualizing the extent of blur in different parts of a given document image. This is created by assigning each visited pixel in set $\mathbb{S}$, the value of quantitative score assigned to the edge profile it belongs to (only if it belongs to one of the unimodal edge profiles). The scores were averaged for pixels which belong to multiple edge profiles and the pixels which do not belong to the set $\mathbb{S}$ were assigned a zero value. The obtained results over couple of images from SOC dataset are illustrated in Figure 4.2. The results clearly demonstrate the capability of the proposed method for local image analysis which can be useful for many important applications like image denoising and enhancement.

The proposed algorithm is also agnostic to font size, font type or language in which the text is written. The algorithm is also not limited to quality assessment of document images with printed text. To demonstrate this aspect, we show the heat map obtained using our algorithm on an image with handwritten text in Hindi language in Figure 4.3. We can observe that the proposed algorithm accurately captures the blur variation inside the camera captured image which further illustrates the generalization ability of the proposed algorithm.

5. CONCLUSION

We have proposed a novel operator and a framework which can be used for the quality estimation of a document image in real time. We have focused on blur detection in this work as it is the prominent source of degradation in document images. Various other kinds of degradations due to document semantics such as broken lines in an image etc. are rarely observed nowadays. Using an extensive experimental evaluation, we have demonstrated that the proposed method out-

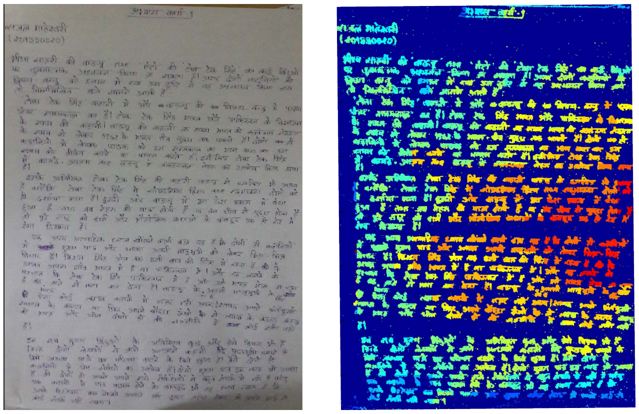


Figure 5: Obtained heat maps using our approach on a handwritten document in Hindi language (Red represents the sharp regions and blue represents blurred regions).

performs both metric and learning based approaches in current state-of-the-art by a significant margin. Furthermore we highlight the ability of our algorithm for complementary application of creating blur maps and demonstrate how it generalizes over varying document types and styles. The real time operation of the proposed algorithm also makes it suitable for many practical applications.

6. REFERENCES

[1] Abbyyfinereader. https://www.abbyy.com/finereader/.

[2] Omnipage. http://www.nuance.com/.

[3] L. R. Blando, J. Kanai, and T. A. Nartker. Prediction of ocr accuracy using simple image features. In *ICDAR*, 1995.

[4] F. Chen, S. Carter, L. Denoue, and J. Kumar. Smartdcap: semi-automatic capture of higher quality document images from a smartphone. In *Proceedings of the 2013 international conference on Intelligent user interfaces*, 2013.

[5] B. Epshtein, E. Ofek, and Y. Wexler. Detecting text in natural scenes with stroke width transform. In *CVPR*, pages 2963–2970, 2010.

[6] R. Ferzli and L. J. Karam. A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb). *IEEE Transactions on Image Processing*, 18(4):717–728, 2009.

[7] R. Hassen, Z. Wang, and M. Salama. No-reference image sharpness assessment based on local phase coherence measurement. In *ICASSP*, 2010.

[8] L. Kang, P. Ye, Y. Li, and D. Doermann. A deep learning approach to document image quality assessment. In *ICIP*, 2014.

[9] J. Kumar, F. Chen, and D. Doermann. Sharpness estimation for document and scene images. In *ICPR*, 2012.

[10] J. Kumar, P. Ye, and D. Doermann. A dataset for quality assessment of camera captured document images. In *International Workshop on Camera-Based Document Analysis and Recognition*, pages 113–125, 2013.

[11] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *Nature*, 521(7553):436–444, 2015.

[12] A. Mittal, A. K. Moorthy, and A. C. Bovik. No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing*, 21(12):4695–4708, 2012.

[13] A. K. Moorthy and A. C. Bovik. Blind image quality assessment: From natural scene statistics to perceptual quality. *IEEE Transactions on Image Processing*, 20(12):3350–3364, 2011.

[14] N. D. Narvekar and L. J. Karam. A no-reference image blur metric based on the cumulative probability of blur detection (cpbd). *IEEE Transactions on Image Processing*, 20(9):2678–2683, 2011.

[15] X. Peng, H. Cao, K. Subramanian, R. Prasad, and P. Natarajan. Automated image quality assessment for camera-captured ocr. In *ICIP*, 2011.

[16] S. V. Rice, F. R. Jenkins, and T. A. Nartker. *The fifth annual test of OCR accuracy*. Information Science Research Institute, 1996.

[17] M. Rusiñol, J. Chazalon, and J.-M. Ogier. Combining focus measure operators to predict ocr accuracy in mobile-captured document images. In *Document Analysis Systems (DAS), 2014 11th IAPR International Workshop on*, pages 181–185, 2014.

[18] J. Shi, L. Xu, and J. Jia. Discriminative blur detection features. In *CVPR*, 2014.

[19] R. Smith. An overview of the tesseract ocr engine. 2007.

[20] P. Ye and D. Doermann. Learning features for predicting ocr accuracy. In *ICPR*, 2012.

[21] P. Ye and D. Doermann. Document image quality assessment: A brief survey. In *ICDAR*, 2013.

[22] P. Ye, J. Kumar, L. Kang, and D. Doermann. Unsupervised feature learning framework for no-reference image quality assessment. In *CVPR*, 2012.

[23] P. Ye, J. Kumar, L. Kang, and D. Doermann. Real-time no-reference image quality assessment based on filter learning. In *CVPR*, 2013.