OCR PERFORMANCE PREDICTION FOR SCENE TEXT IMAGES WITH DIQA METHODS: A COMPARISON BASED STUDY

Author(s): Sabbir Ahmed Sibli, Nuray Jannat, Shihab Sharar, Annajiat Alim Rasel

Presentation By:
Sabbir Ahmed Sibli, 20266027
Nuray Jannat, 20266028
M.Sc. in CSE, BRAC University

TABLE OF CONTENTS

BACKGROUND STUDY

- Optical Character Recognition opens a new window for NLP as it is undoubtedly a better way to extract information from the images with texts.
- But quality of images with texts plays a vital role in character recognition optically. Because, bad image hinders perfect text extraction from it.
- In term of image quality evaluation, Image Quality Assessment (IQA) and Document Image Quality Assessment (DIQA) are different. Because IQA focuses on quality of whole image while DIQA focuses on quality of the texts in an image.
- There are a number of DIQA methods to determine OCR accuracy or eligibility of images with texts.
- So, a comparison among different DIQA is needed for finding the better approach to evaluate the usability of images with texts.

LITERATURE SURVEY

- Assessing the quality of degraded text images using unsupervised DIQA approaches (Peng Ye et. al., 2012).
- Character recognition with No-reference Document Image Quality Assessment Based on Character Gradient (Hongyu Li et. al., 2018).
- (Peng Ye et. al. 2013) showed a survey on Document Image Quality Assessment including IQA, DIQA etc.
- OCR Accuracy Prediction Method Based on Blur Estimation of an Image (Van-Cuong Kieu et. al., 2016).
- Document Image Quality Assessment Method Based on DIQA: Hast Derivations (Alireza Alaei, 2019).
- A dataset has been prepared of camera captured document images containing varying levels of focal-blur introduced manually during capture (Jayant Kumar et. al., 2013).

PROPOSED MODEL

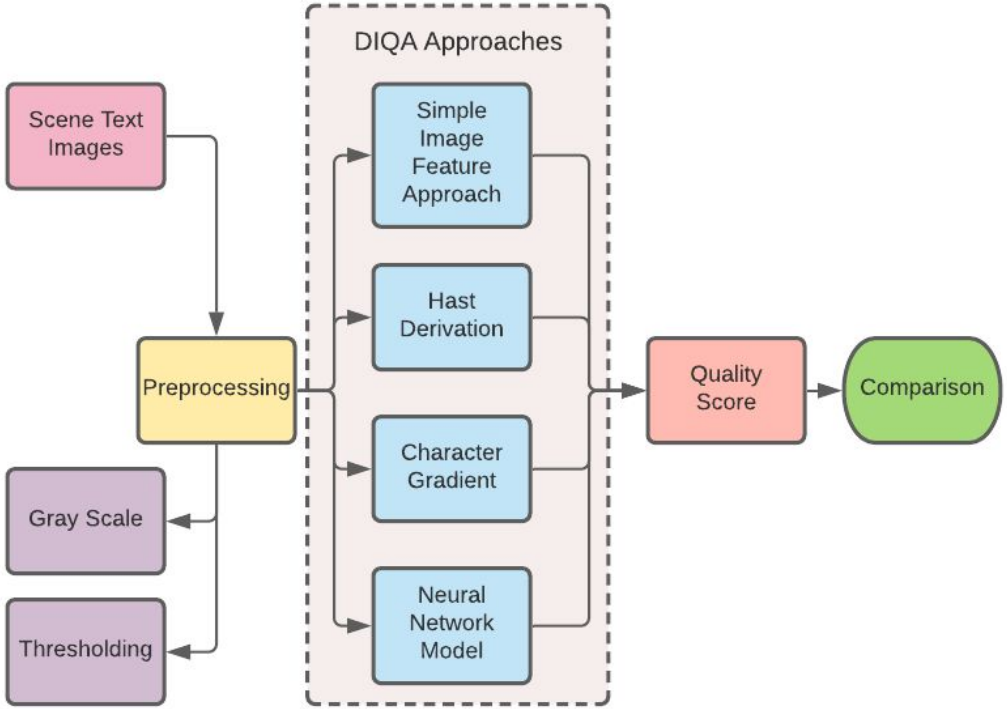


Fig. Workflow of our Proposed Model

DATASET OVERVIEW

- "The Street View House Numbers (SVHN) Dataset" by Stanford University has been used as the dataset of this research.

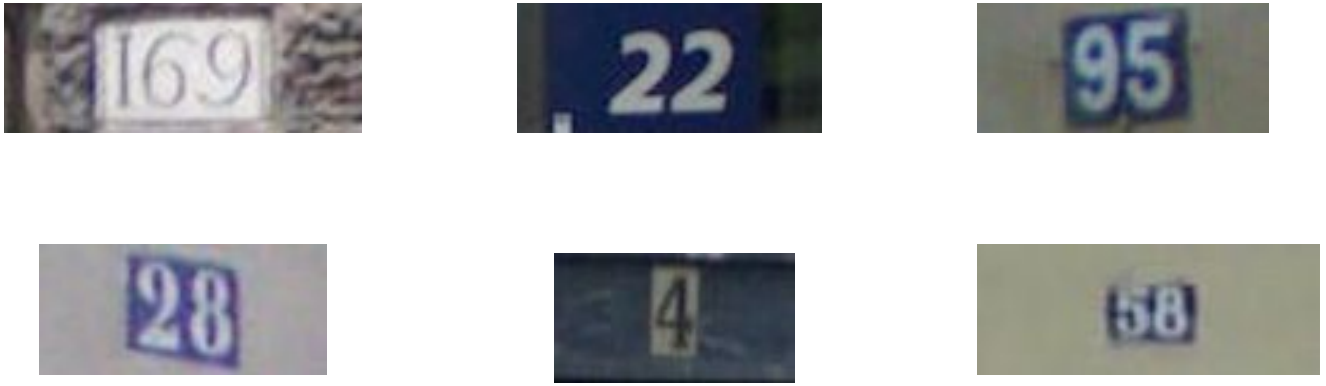


Fig: Few images from dataset

DATA PREPROCESSING

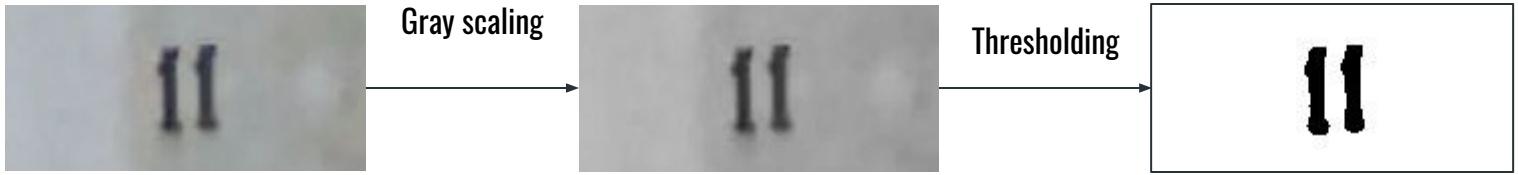


Fig: Preprocessing flow

SIMPLE IMAGE FEATURES

- There are a number of simple image features which can be used for finding:
  - Touching characters
  - Broken characters
  - White letters on black background

- Some simple image features are:
  - White Speckle Factor (WSF): determines touching characters
    IF WSF >= 0.1 THEN poor
  - Broken Character Factor (BCF): determines broken characters
    IF BCF >= 0.7 THEN poor
  - Number of white connected components and their sizes: detects white letters on black background
    IF (max(WhiteWidth,WhiteHeight) >= 30 pixels)

HAST DERIVATION

Fig: Hast Derivation Working Flow

CHARACTER GRADIENT

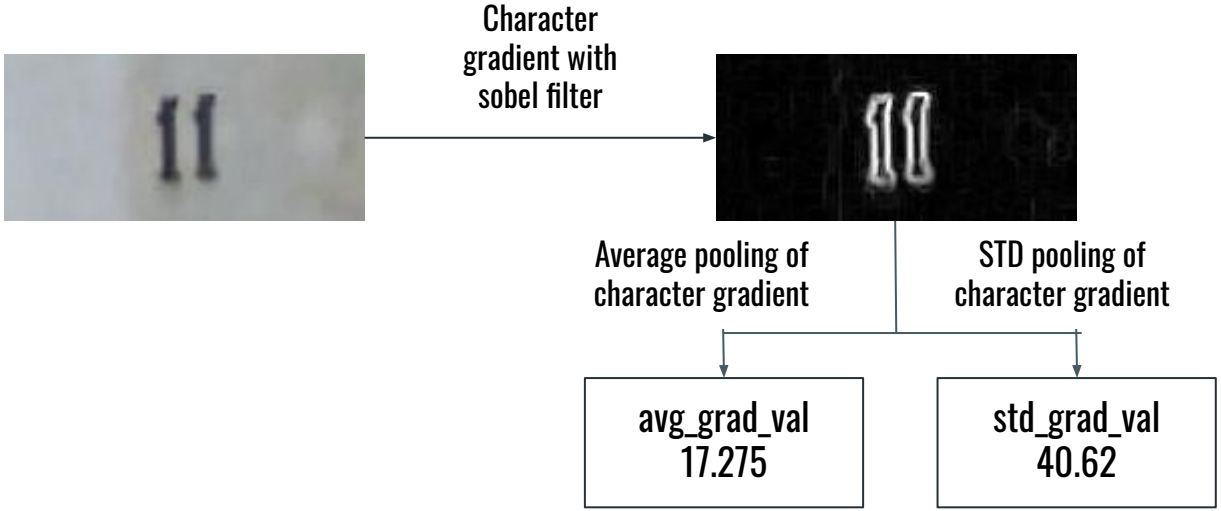
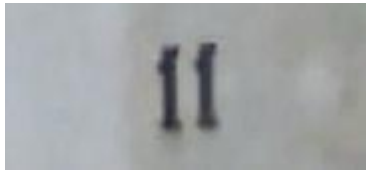
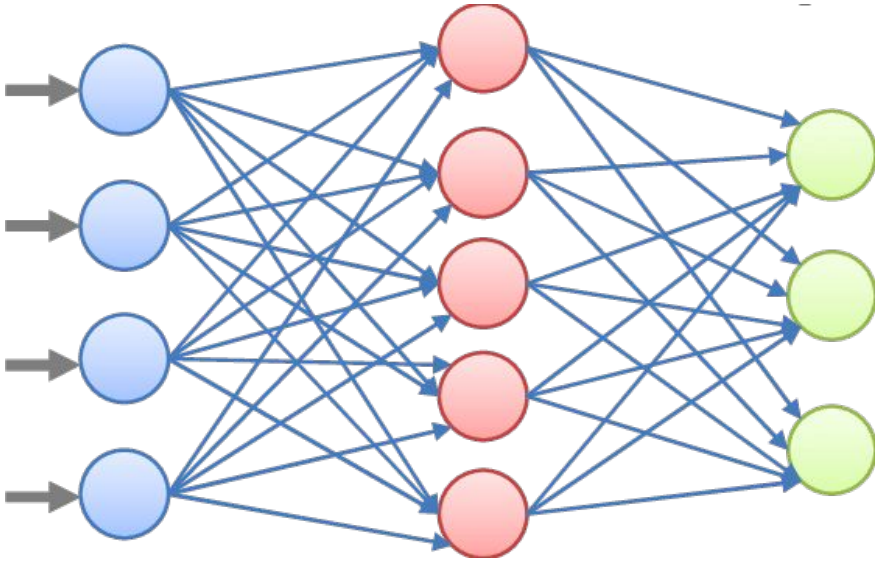


Fig: Character gradient working flow

QUALITY SCORE PREDICTION WITH NEURAL NETWORK

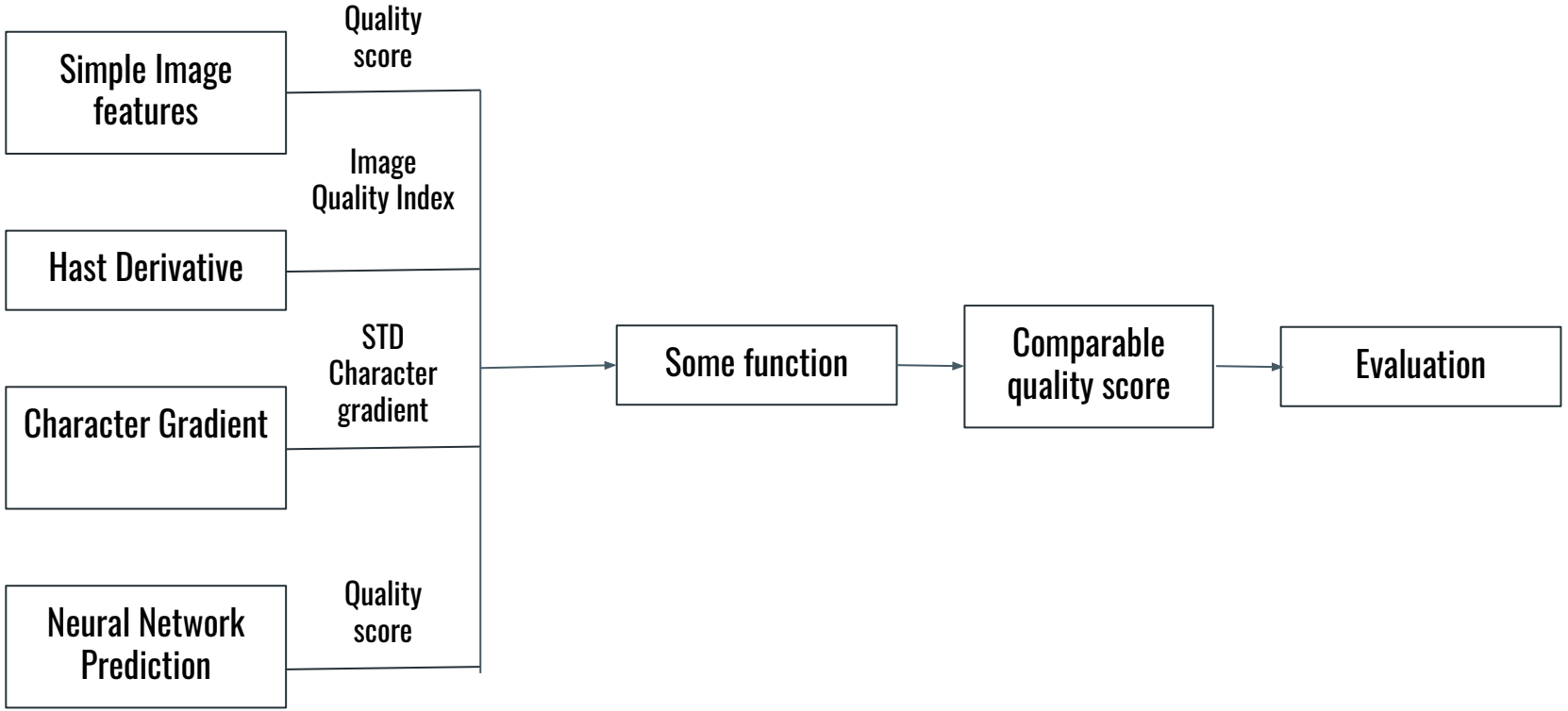




Quality score 0.78

Fig: Quality Score Prediction flow working flow

COMPARISON AMONG THE DIQA METHODS



CONCLUSION

- There are number of DIQA methods. But it is difficult to determine which one is perfect to use.
- Our work shows a comparison among various DIQA methods using a specific dataset. But it also helps to find a better working ability for other dataset's particular domain as well.
- For different dataset, these methods can be compared to be more sure about their ability.

“

Every great work of art, starts with an act of imagination.

— JONAH LEHRER

REFERENCES

1. Ye, P. and Doermann, D., 2012, November. Learning features for predicting OCR accuracy. In *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)* (pp. 3204-3207). IEEE.
2. Li, H., Zhu, F. and Qiu, J., 2018, August. CG-DIQA: No-reference document image quality assessment based on character gradient. In *2018 24th International Conference on Pattern Recognition (ICPR)* (pp. 3622-3626). IEEE.
3. Ye, P. and Doermann, D., 2013, August. Document image quality assessment: A brief survey. In *2013 12th International Conference on Document Analysis and Recognition* (pp. 723-727). IEEE.
4. Kieu, V.C., Cloppet, F. and Vincent, N., 2016, April. OCR accuracy prediction method based on blur estimation. In *2016 12th IAPR Workshop on Document Analysis Systems (DAS)* (pp. 317-322). IEEE.
5. Alaei, A., 2019, September. A new document image quality assessment method based on hast derivations. In *2019 International Conference on Document Analysis and Recognition (ICDAR)* (pp. 1244-1249). IEEE.
6. J. Kumar, P. Ye, and D. Doermann, “A Dataset for Quality Assessment of Camera Captured Document Images,” in Proceedings of the 2013 Camera-Based Document Analysis and Recognition, ser. CBDAR’13, Washington DC, USA, 2013, pp. 113–125.
7. Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B. and Ng, A.Y., 2011. Reading digits in natural images with unsupervised feature learning.