Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 133–142

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

Quality Assessment Model for Handwritten Photo Document Images

N Shobha Rani [a,*], Akshatha K.A [a], Koushik K.S [b]

[a]*Department of Computer Science, Amrita Vishwa Vidyapeetham, Mysuru 570026, India.*
[b]*Department of Computer Science, Amrita Vishwa Vidyapeetham, Mysuru 570026, India.*

Abstract

The photo document image's quality determines whether it has the potential to be used for information extraction. Document Image Quality Assessment (DIQA) is a difficult task since it is complicated to train a system to have a complete human-like vision and it might be tiresome and time-consuming to manually evaluate the quality of many document images. The educational system became online during the pandemic age, which led to the digital submission of exams and homework. Quality evaluation has gained more prominence recently as digitization has started to take precedence in some fields. In order to evaluate the document image quality of handwritten document images at the page level while taking into account the overall visual aesthetic, this research suggests a transfer-learning and classification-based model. ResNet50 exceeded all of the pre-trained models we tested and examined, with an accuracy of 91.00 % on validation sets and 83 % on test sets.

Keywords: Document Image Quality Assessment (DIQA) ; ResNet50 ; Transfer Learning ; Deep-learning.

1. Introduction

In today's world, the use of computers or digital media has skyrocketed in almost every field. To develop a paperless environment, it is necessary to digitize documents such as receipts, bills, assignments, and answer sheets. Scanners, digital cameras, mobile phone cameras, and other devices can be used to digitize data. A document is a depiction of a

* N Shobha Rani, Tel.: 0821-234 3479;
E-mail address: n_shobharani@my.amrita.edu

10.1016/j.procs.2022.12.409

piece of written, printed, or electronic substance that serves as an official transcript that provides proof or data. Printed documents, handwritten documents, and mixed materials are among the existent documents. A document can go through multiple phases of scanning, printing, transmission, compression, and decompression, making a document degrade, and as a result, a single degradation may occur frequently, or combinations of multiple degradations may appear in unexpected ways [1]. Other steps that have an impact on document image quality due to physical or environmental conditions, such as the amount and direction of lighting used to capture the document. Handwriting is a crucial skill for acquiring and transmitting information and is linked to decipherability and kinematics [2], while legibility and kinematics are linked to document quality [3]. Current quality evaluation systems rely almost entirely on manual labour, which has several significant drawbacks, including varying assessment norms from one person to the next, low proficiency, and high labour costs. Tasks like document image classification, recognition, and detection of characters, words, etc. generally utilize document image processing, but handwritten document quality ratings do not [3]. The evaluation can be done at several levels, including page, phrase, word, and character. Existing assessment or analysis models mainly focus on character-level evaluations that compare symbols to a reference model, with the majority of systems proposing an analysis based on five metrics: direction, kinematics, order, form, and location relative to reference lines [4]. Some of the examples of page-level degradation include background noise like salt and pepper noise, ink bleeds, cluttering, and structural deformations due to skew, warping, folds, and curling [1]. DIQA usually takes one of the two routes: Traditional approach or a modern approach like Machine learning / Deep learning.

Traditional image processing approaches have typically been used to examine the image, which contains metrics or a reference to which the image may be compared. Image acquisition, pre-processing, edge detection, and feature extraction are some of the most widely used procedures in character or word assessment. The hardcopy or paper is scanned using a scanner, or an image of the document is captured using any imaging equipment such as a digital camera or a mobile phone camera. The acquired image is considered raw data that must be processed in order to improve the image and remove noise or distortions that could damage the results. The images are resized in this phase to standardize the dataset, eliminating the need for costly recurrent symbol alignment and creating a common ground for the images with the goal of simplifying the analysis process by organizing the data [5] [6]. Edge Detection is a method of segmenting an image into zones of discontinuity that makes it easier to spot features in an image that have a large shift in grey level. This minimizes the amount of data in an image while maintaining its structural qualities. The recognition of handwritten characters is greatly aided by a rich feature set. A set of good features is used to represent a handwriting text in feature extraction [6]. By lowering the number of features and considering relevant features, training and testing time can be lowered [7].

When compared to traditional methods that rely on a single feature, the deep learning method inspired by the human nervous system has shown to improve performance in any field while also allowing for the investigation of a set of features that may not be readily apparent through visual inspection [8]. Deep nature has the advantage of incorporating features from several levels, whereas typical state-of-the-art techniques focus on either low-level or high-level features. Despite the fact that deep networks produce excellent results and have the benefit of reducing training costs and time, there is still a risk of overfitting and increased training error [9]. In this paper section 2 discusses existing literature, section 3 describes the proposed method, next section 4 provides a brief description of the experiments conducted, and then section 5 discusses the results obtained followed by conclusion in section 6.

2. Related Work

The DIQA methods can be addressed as learning-based and metric-based assessment method, as well as reference and no-reference approaches based on feature extraction methods. Here we discuss DIQA in two categories: Traditional method and Deep Learning-based methods [10].

2.1. Works based on traditional methods

Bouillon et al., [11] present work for evaluating children's handwriting (symbols, letters in any geometric form) that uses the generative and discriminative capabilities of fuzzy inference systems to evaluate handwritten symbols concerning geometric features like shape and the direction the used feature set on three specific feature sets to assess the morphology, sequence, and direction of the symbols where it can be observed that the out is highly dependent on the input provided. While the paper by Akouaydi et al., [4] deals with the handwriting quality analysis of Arabic letters employing the Beta elliptical model for segmentation of characters to access them closely and the Cartesian Fourier Descriptor model to find the shape or boundary of the character. The inputs are taken through a tablet hence creating a scenario for pre-processing and the system is trained with 20 samples of each correctly written letter where the results might be questionable due to limitation of using a small dataset. Later, Simonnet et al., [12] presented a system in which handwritten words are collected as on-line signals on electronic touchscreens with styluses and corrected with the EVOLVE classifier. The dataset is focused on the French language, and some examples of this collection contain missing letters. As it is primarily made up of up-strokes, the paper tests a variety of segmentation methods and hypotheses. Another approach presented by Simonnet et al. [13] focuses on the development of a cursive writing analyser that provides global feedback based on global, directional, geometrical, and order features while comparing inter-class and intra-class scores with a model that acts as a reference. The multi-criteria classifier incorporates the findings of the classifiers of the above-mentioned features to produce a final result that considers all of these factors. The words in this method are evaluated on IntuiScript, a digital notebook while incorporating the stable features of block letters to improve the overall approach [13]. Next, Li et al., [10] propose a DIQA based on a no reference method MSER along with employing the accuracy obtained by the OCR as a criterion to measure the quality with the assumption of patches containing critical features which may not be entirely consistent in terms of human perception, given the quality of the entire document can weigh in on aesthetic perception, and the dataset in question consists of characters of uniform shape.

Subsequently, Kulesh et al. [14] proposed a model based on low-level features such as the aspect ratio of height to width, zero-crossing distributions, and width distributions across the height of the character, and mapped them into high-level feature vectors. To evaluate and score each handwritten letter, this paper combines the best of an artificial neural network and an expert system but it is complex to understand and implement. In a work reported by Nayef et al., [15] have employed no reference and metric-based methods of quality assessment that take sharpness quality metric and character quality metric where sharpness is estimated by LPC Sharpness Index using Log-Gabor filters of $M \times N$ scales and character quality is assessed by considering the black and white speckle-noise along with overlapping or touching characters, while mostly focusing on distortion specific assessment on scanned and mobile captured images of bills, receipts, etc., and also 175 images from the Tobacco dataset while comparing with human visual perception. The study by Alaei [16] proposes a new full-reference DIQA approach based on Hast derivations, where a similarity map is constructed on both reference and distorted pictures. The quality of the distorted document image is then assessed using average pooling on the ITESOFT and LIVE datasets, with the results indicating that the second-order Hast derivation performed better on document images, while the first order Hast derivation performed better on natural scene images [16].

2.2. Work based on Deep-Learning Methods

Learning-based DIQA methods extract discriminant characteristics for various types of document degradations using learning techniques such as deep learning [17], [18]. Gao et al., [3] introduced and evaluated a CNN-based English Handwriting Evaluation Algorithm to existing classification methods. The dataset is divided into two types of handwriting: award-winning and non-award-winning. Because of the dataset imbalance, the ResNet-18 model was trained with weights, with softmax as the loss function and Stochastic Gradient Descent (SGD). Kang et al., [17] use a learning-based approach to assess the quality of the SOC dataset, which contains 175 color images, and the Newspaper dataset, which contains 571 grayscale images, using a CNN with 2 convolution layers, location blind maxmin pooling, and Rectified Linear Units in the fully connected (ReLU) layers, and the patch scores are averaged over the image to obtain the document score. However, because the images are largely printed document images, the dataset does not represent much of a problem in terms of homogeneity. Although splitting the image into patches may provide additional samples for the CNN, this cannot be compared to or associated with human visual perception. In the similar approach proposed by Lu et al., [19], a Deep CNN with 3 fully-connected layer is trained and fine-tuned, all while using OCR as a quality descriptor.

The DIQA strategy proposed by Peng et al., [18], calculates the target quality score using sparse code learning, in which the quality of the photograph is encoded into the codebook using a pooling approach that vectorizes the full document image, and each patch from a training image can be expressed by a combination of code words to build a model that can be used to map images to OCR confidences, and linear regression training is performed. Alaei et al., [20] also consider a blind approach trained on the ITESOFT dataset, segments the image into patches to further create a BoW (Bag of Words), based on which the features extracted from the patch in the testing phase are assigned to the cluster that is subjected to average pooling where the quality of the image is measured IQA is between 0 to 1, where 0 indicates a low-quality image and an IQA that is closer to 1 indicates a better quality image. Li et al., [21] propose an RNN-based DIQA model that employs a convolutional layer for feature extraction from key sections of images determined using a spatial glimpse, and trains the model into parts using reinforcement learning and SGD. On two datasets, Smartdoc-QA and SOC, the proposed system was trained and tested. The system did not perform well on Smartdoc-QA since it was unable to integrate degradations such as shadows.

2.3. Motivation

- The quality of an image can be assessed in a variety of methods, including a step-by-step classical methodology or newer technologies such as deep learning. The state-of-art works mostly concentrate on the degradation on a character level, word level while the main focus of this paper is to address the degradation on a page level.
- There is no readily available benchmarking dataset that includes varieties of degradation and degradation in real-time.

2.4. Comparative Analysis

Table 1: Comparative analysis of reviewed papers.

Author	Problem Addressed	Method
Bouillon et al., [11]	Character level evaluation of handwritten documents dependent on provided input.	Fuzzy Inference.
Akouaydi et al., [4]	Quality analysis of Arabic letters in Handwritten samples with limited training samples.	Cartesian Fourier and Descriptor model.
Simonnet et al., [12]	Quality analysis of French letters on Handwritten samples.	EVOLVE classifier.
Simonnet et al., [13]	Quality analysis of cursive writing letters on Handwritten samples on digital notebook that exclude challenges like degradation.	Multi-criteria classifier based on various features
Nayef et al., [14]	Distortion specific assessment on scanned & mobile captured images of bills, receipts, etc.	Log-Gabor filters.
Gao et al., [3]	Handwriting sample classification with a model built from scratch that is complex and time-consuming.	ResNet-18 model.
Kang et al., [17]	Quality analysis of printed documents like that of the SOC dataset.	CNN
Alaei et al., [20]	Analysing document images as patches makes comparison by visual inspection unfair.	BoW, CNN

Peng et al., [18]	Document image quality analysis based on codebook encoding that works based on the degree of similarity or generalization.	Sparse code learning
Li et al., [21]	Mixed document image quality assessment with RNN is prone to diminishing or exploding gradient.	RNN

3. Methodology

3.1. Proposed System

This study proposes a classification-based transfer learning approach for DIQA where the document is classified into one of the classes 1 to 5 based on the page-level quality. Performance tuning is carried out using the hyper-parameters such as number of epochs, batch size and sampling size of the datasets and number of linear units with normalization [22]. Despite the improved results, building and training a deep-learning model from the ground up is often complex and time-consuming, which can be solved using transfer learning, which is a deep-learning method for improving learning in a new task that attempts to change traditionally isolated tasks by developing methods to transmit the previously learned knowledge from a single or multiple source tasks and then use it for the improvement of learning on a similar task [23]. The transfer of expertise gained by a machine learning model in one domain to another related area is known as transfer learning [24]. By removing the requirement of data independence that is identically distributed for testing, we can apply transfer learning to alleviate the problem of insufficient training data [25].

The basic structure of a deep learning model can be broken down into the Input layer, Convolutional layer, ReLU (Rectified Linear Unit), Fully Connected Layer, and Softmax layer where the proposed model retains the original architecture of the pre-trained model beginning with a convolutional layer with a 7×7 kernel size and 64 filters, with a modification in the amount of movement over the image set to move by 2 units or pixels at a time, sequentially followed by Batch Normalization and ReLU activation, a 3×3 Max Pooling layer, four blocks of ResNet, an Average-Pooling layer, a fully connected layer, and a softmax layer [26]. The model is built sequentially, with the pre-trained model added first, followed by a flattening layer to convert the 2D arrays in pooled feature maps into a single 1D or linear continuous vector. We build a Deep CNN with a fully connected layer with softmax activation using the pre-trained ResNet50 model with 'imagenet' weights. The layers are initially frozen, then trained while fine-tuning, which saves space and time by ensuring the model does not repeat the weight-learning process without any dropout.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
resnet50 (Functional)        (None, 2048)              23587712
_________________________________________________________________
flatten (Flatten)            (None, 2048)              0
_________________________________________________________________
dense (Dense)                (None, 512)               1049088
_________________________________________________________________
dense_1 (Dense)              (None, 5)                 2565
=================================================================
Total params: 24,639,365
Trainable params: 24,586,245
Non-trainable params: 53,120
```

Fig. 1. Summary of the proposed model.

3.2. Workflow the Model

The images in the dataset are created and collected from different environments resulting in diverse image sizes. The images are augmented, resizing the images into images of 224 × 224. The 2-layer block in the 34-layer net with a 3-layer bottleneck block results in a 50-layer ResNet50 model. The pre-trained model is loaded from the Keras machine learning library that is added to a Sequential model.

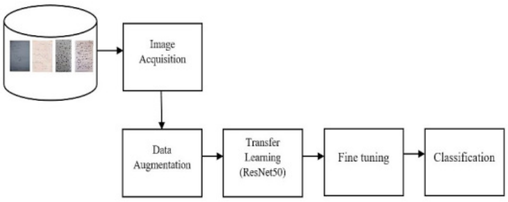

Fig. 2. Workflow of the proposed model.

The Adamax optimizer, a version of the Adam optimizer based on infinity form [27], is used in the model with a batch size of 32, a learning rate of 0.0001, and a decay rate of 0.00001, and 'categorical_crossentropy' as the loss function. Hyper parameters can make a big difference in the training process and can also help to fit a model better. Callbacks such as early stopping and checkpoints have been designed to monitor the model and focus on validation loss to avoid the model from overfitting. The EarlyStopping function is set to monitor validation loss, and the mode is set to minimal with patience and verbose set to 1 so that the model stops training when the value of validation loss does not improve.

3.3. Fine-tuning

Fine-tuning is the process of updating weights during back-propagation to alter the mean and variance of the model's top layers, which may differ from the mean and variance of our dataset [8]. The layers of the model imported from the Keras library are initially frozen in the training stage causing the weights to not be updated, making the layers of the model non-trainable. After the training phase, all the layers of the model are unfrozen and the model has trained again with a lower learning rate. To fine-tune the model, hyper parameters that aid in the regulation of the learning process is modified, where the learning rate is set to 0.00001, while the decay rate is set to 0.000001, with the batch size and epochs remaining constant.

4. Experiments

4.1. Tools used

The model is trained in Keras Tensorflow (2.6.2) environment on Kaggle Notebook, which is a cloud computational environment on a system with Windows 10 with a RAM of GPU. Some of the libraries used include Tensorflow, OpenCV, PIL, Matplotlib and Numpy.

4.2. Dataset

As the DIQA dataset is not readily accessible online, a customized dataset must be constructed. Images of handwritten documents with various forms of aberrations and quality are included in the dataset. The photographs are acquired with a cell phone camera, which leaves them vulnerable to degradation owing to lighting, blur, or compression quality loss. Due to the varying ambient conditions, the distance and angle at which the image is captured, and the device used to acquire the image, the custom dataset presents a range of obstacles.

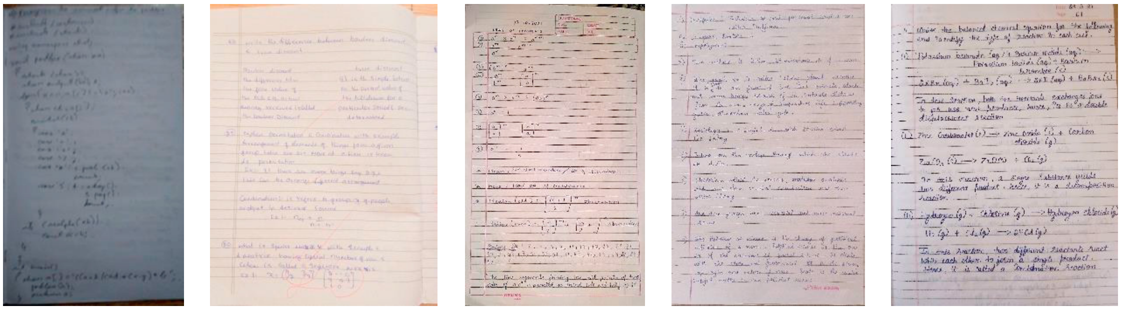

Fig. 3. (a) Class 1 image; (b) Class 2 image; (c) Class 3 image; (d) Class 4 image; (e) Class 5 image.

The dataset contains photos of student notes, assignments, and tests from various educational institutions. The dataset is categorized into 5 groups, with 1 denoting a document image with a higher level of degradation and 5 denoting a document image with no or a lower level of degradation. A subset of the training dataset is used for validation, and the dataset is split into training and test data. To train the model, we used our custom dataset and a few samples from the CSAFE dataset [28] and the formsl-Z IAM dataset [29]. Photos from the custom dataset as well as a Denoising dataset from Kaggle that contains images of documents with degradations like stains and textures were used to test the model The dataset is split into 3 categories: training, testing, and validation in the ratio of 80: 10: 10. All the images in the dataset are of varied sizes and hence have to be normalized and the images are labelled as one of the five classes. It is difficult to even create a rather smaller dataset of 10,000 images. Some of the images in the dataset have been altered while creating the dataset as most of the images are available to create images for different labels. We have tried to train the model by training by splitting the dataset into different ratios.

4.3. Models

Training a model from scratch can be complex and time-consuming. To overcome this issue, a model can be trained on a readily available or custom dataset using the existing pre-trained models. We have experimented with pre-trained ResNet50, ResNet50 V2, ResNet101, ResNet152, Xception and VGG16 models having 50, 50, 101, 152, 71 and 16 dense layers respectively. The hyper-parameters such as rate of learning and decay, number of epochs, and batch size significantly contribute to the training of a model. For the current dataset, the ResNet50 yielded the best results. It is considered most suitable to use the method of transfer learning when the dataset is not extensively small or without significant similarities.

5. Results

The proposed model's performance is evaluated using the following metrics: accuracy, loss, and confusion matrix. After fine-tuning, the model's accuracy shows to have improved slightly. The performance of Model A, Model B, Model C, Model D, Model E, and Model F, which use pre-trained ResNet50, ResNet50V2, ResNet101, ResNet152, Xception, and VGG16 models, is shown in the graphs below. The Adamax optimizer was used to fine-tune all of the models with the same learning rate of 0.00001 and decay rate of 0.000001. The total number of epochs considered for training, as well as the Early Stopping function with patience = 1. Model B terminated early in the training phase, at the third epoch, with validation accuracy and validation loss of 0.5594 and 2.6866 but an improvement in the performance of Model B improved after fine-tuning. Model C learns throughout 15 epochs of training and fine-tuning, where we can notice from Table 2 and Table 3 that the validation accuracy has improved from 0.8717 to 0.9067, while the validation loss reduced from 0.3425 to 0.2428. Model D ends the training phase early, at the tenth epoch, with validation accuracy and validation loss of 0.8483 and 0.3765, respectively, and completes 15 epochs in the fine-tuning phase, yielding an improved validation accuracy and validation loss of 0.9111 and 0.2239

Table 2: Performance of models after training

Name of the Model	Accuracy in Training	Loss in Training	Accuracy in Validation	Loss in Validation
Model A (ResNet50 V2)	0.6144	2.7525	0.5594	2.6866
Model B (ResNet101)	0.8942	0.3085	0.8717	0.3425
Model C (ResNet152)	0.8865	0.3246	0.8483	0.3765
Model D (Xception)	0.9309	0.6591	0.5544	1.2224
Model E (VGG16)	0.8942	0.3085	0.8717	0.3425

Table 3: Performance of models after fine-tuning

Name of the Model	Accuracy in Fine-tuning	Loss in Fine-tuning	Accuracy in Validation	Loss in Validation
Model A (ResNet50 V2)	0.9414	0.2151	0.8511	0.3673
Model B (ResNet101)	0.9814	0.0877	0.9067	0.3428
Model C (ResNet152)	0.9888	0.0666	0.9111	0.2239
Model D (Xception)	0.9732	0.0785	0.9206	0.2277
Model E (VGG16)	0.9817	0.0671	0.8922	0.3039

Model E terminates early in both the training and fine-tuning phase, at the fifth and tenth epoch respectively. It can be observed from Table 2 that there is a high improvement in validation accuracy but the model misclassified more images compared to any other model. Model F learning overall 15 epochs, with validation accuracy and validation loss of 0.8942 and 0.3085, respectively. With small increases in validation accuracy and loss, it stops early in the fine-tuning phase at the eighth epoch.

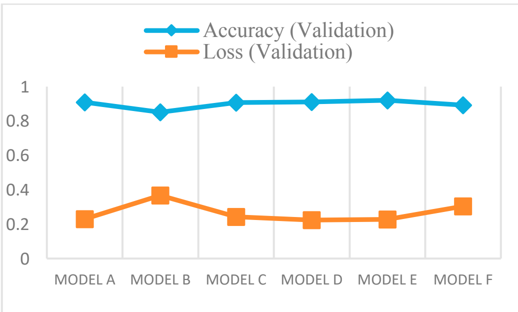

Fig. 4. Comparison of Models performance after fine tuning.

It can be observed from Figure 4, that Model A (ResNet50) and Model D (ResNet152) have better accuracy compared to the other models. ResNet152 has slightly higher accuracy compared to ResNet50 but the degradation caused while training should also be considered when choosing the most suitable model. When the models are tested side by side, ResNet152 misclassified more images than ResNet50.

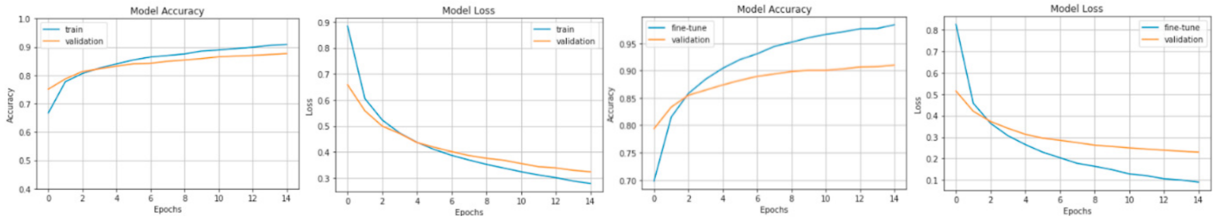

Fig. 5. (a) ResNet50 - Accuracy after training (b) ResNet50 - Loss after training (c) ResNet50 - Accuracy after fine-tuning (d) ResNet50 - Loss after fine-tuning.

When compared to the other models in the study, ResNet50 outperforms them all. The performance of a model cannot be only determined by its accuracy, as the performance of a model can be questioned whether the accuracy is very low or very high, as models with higher accuracy are more likely to be prone to overfitting, resulting in inaccurate results.

Table 4: Performance of Model A (resnet50 model) after training and fine-tuning

Metric	Outcome in training phase	Outcome in fine-tuning phase
Accuracy (Training)	0.9088	0.9839
Loss (Training)	0.2785	0.0903
Accuracy (Validation)	0.8767	0.9100
Loss(Validation)	0.3233	0.2298

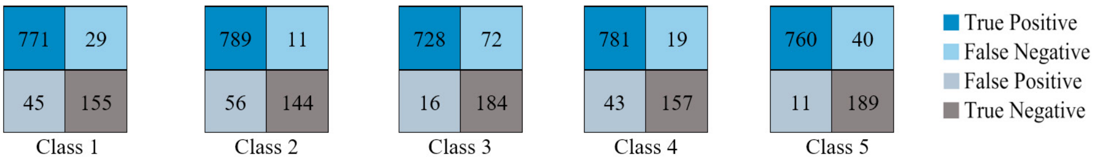

Fig. 6. Confusion matrix for each label [1 to 5] depicting True-positive, True-negative, False-positive, False-negative.

The confusion matrix is a statistic that represents the model's overall prediction summary, allowing us to see how many images were correctly predicted and how many were incorrectly predicted. Figure 6 shows the TP, FN, FP, and TN values for each class. The confusion matrix can be easily plotted with sklearn multilabel_confusion_matrix.

6. Conclusion

The classification-based DIQA proposed in this work employs transfer learning along with GPU that aids in building better systems. DIQA is a challenging task that may require tedious brute force experimentation, evaluation and comparison of results which are subjective in nature. The system achieves a considerable level of accuracy that can be improved with further fine-tuning and also experimenting with newer models and gives results closer to human perception as the performance of DIQA is judged through visual inspection. The performance of each model is greatly affected the different hyper-parameters where the difference in results may range from a few points to a vast percentage. In the future, we plan on extending the model to assess the quality at the word level and character level while improving the model to perform on a higher scale with handwritten documents of various scripts, participants and broader niche which gives it the advantage the improvement each time the model is retrained.

References

[1] Ye P, Doermann D. Document image quality assessment: A brief survey. In2013 12th International Conference on Document Analysis and Recognition 2013 Aug 25 (pp. 723-727). IEEE.
[2] Dinehart LH. Handwriting in early childhood education: Current research and future implications. Journal of Early Childhood Literacy. 2015 Mar;15(1):97-118.
[3] Gao Y, Yu R, Duan X. An English Handwriting Evaluation Algorithm Based on CNNs. In2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM) 2019 Oct 16 (pp. 18-21). IEEE.
[4] Akouaydi H, Hamdi Y, Boubaker H, Zaied M, Alaya Cheikh F, Alimi A. Children's Online Handwriting Quality Analysis.
[5] Ptucha R, Such FP, Pillai S, Brockler F, Singh V, Hutkowski P. Intelligent character recognition using fully convolutional neural networks. Pattern recognition. 2019 Apr 1;88:604-13.
[6] Katiyar G, Katiyar A, Mehfuz S. Off-line handwritten character recognition system using support vector machine. American Journal of Neural Networks and Applications. 2017 Dec;3(2):22-8.
[7] Ghadekar P, Ingole S, Sonone D. Handwritten digit and letter recognition using hybrid dwt-dct with knn and svm classifier. In2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) 2018 Aug 16 (pp. 1-6). IEEE
[8] Reddy AS, Juliet DS. Transfer learning with ResNet-50 for malaria cell-image classification. In2019 International Conference on Communication and Signal Processing (ICCSP) 2019 Apr 4 (pp. 0945-0949). IEEE.
[9] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. InProceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 770-778).
[10] Li H, Zhu F, Qiu J. CG-DIQA: No-reference document image quality assessment based on character gradient. In2018 24th International Conference on Pattern Recognition (ICPR) 2018 Aug 20 (pp. 3622-3626). IEEE.
[11] Bouillon M, Anquetil E. Handwriting analysis with online fuzzy models. In17th Biennal Conference of the International Graphonomics Society (IGS) 2015 Jun 21.
[12] Simonnet D, Girard N, Anquetil E, Renault M, Thomas S. Evaluation of children cursive handwritten words for e-education. Pattern Recognition Letters. 2019 Apr 15;121:133-9.
[13] Simonnet D, Anquetil E. Handwriting Quality Analysis of Block Letters and Cursive Words. Handwriting Today, Journal of the National Handwriting Association. 2016 Dec(15):15-21.
[14] Kulesh V, Schaffer K, Sethi I, Schwartz M. Handwriting quality evaluation. InInternational Conference on Advances in Pattern Recognition 2001 Mar 11 (pp. 157-165). Springer, Berlin, Heidelberg.
[15] Nayef N, Ogier JM. Metric-based no-reference quality assessment of heterogeneous document images. InDocument Recognition and Retrieval XXII 2015 Feb 8 (Vol. 9402, p. 94020L). International Society for Optics and Photonics.
[16] Alaei A. A new document image quality assessment method based on hast derivations. In2019 International Conference on Document Analysis and Recognition (ICDAR) 2019 Sep 20 (pp. 1244-1249). IEEE.
[17] Kang L, Ye P, Li Y, Doermann D. A deep learning approach to document image quality assessment. In2014 IEEE International Conference on Image Processing (ICIP) 2014 Oct 27 (pp. 2570-2574). IEEE.
[18] Peng X, Cao H, Natarajan P. Document image quality assessment using discriminative sparse representation. In2016 12th IAPR Workshop on Document Analysis Systems (DAS) 2016 Apr 11 (pp. 227-232). IEEE.
[19] Lu T, Dooms A. A deep transfer learning approach to document image quality assessment. In2019 International Conference on Document Analysis and Recognition (ICDAR) 2019 Sep 20 (pp. 1372-1377). IEEE.
[20] Alaei A, Conte D, Martineau M, Raveaux R. Blind document image quality prediction based on modification of quality aware clustering method integrating a patch selection strategy. Expert Systems with Applications. 2018 Oct 15;108:183-92.
[21] Li P, Peng L, Cai J, Ding X, Ge S. Attention based RNN model for document image quality assessment. In2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR) 2017 Nov 9 (Vol. 1, pp. 819-825). IEEE.
[22] Bengio Y, LeCun Y. Scaling learning algorithms towards AI. Large-scale kernel machines. 2007 Aug;34(5):1-41.
[23] Olivas ES, Guerrero JD, Martinez-Sober M, Magdalena-Benedito JR, Serrano L, editors. Handbook of research on machine learning applications and trends: Algorithms, methods, and techniques: Algorithms, methods, and techniques. IGI global; 2009 Aug 31.
[24] Pan SJ, Yang Q. A survey on transfer learning. IEEE Transactions on knowledge and data engineering. 2009 Oct 16;22(10):1345-59.
[25] Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C. A survey on deep transfer learning. InInternational conference on artificial neural networks 2018 Oct 4 (pp. 270-279). Springer, Cham.
[26] Krishna ST, Kalluri HK. Deep learning and transfer learning approaches for image classification. International Journal of Recent Technology and Engineering (IJRTE). 2019 Feb;7(5S4):427-32.
[27] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. 2014 Dec 22.
[28] Crawford A, Ray A, Carriquiry A. A database of handwriting samples for applications in forensic statistics. Data in brief. 2020 Feb 1;28:105059.
[29] Marti UV, Bunke H. The IAM-database: an English sentence database for offline handwriting recognition. International Journal on Document Analysis and Recognition. 2002 Nov;5(1):39-46.
[30] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J. Scikit-learn: Machine learning in Python. the Journal of machine Learning research. 2011 Nov 1;12:2825-30.
[31] Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M, Ghemawat S, Irving G, Isard M, Kudlur M. {TensorFlow}: A System for {Large-Scale} Machine Learning. In12th USENIX symposium on operating systems design and implementation (OSDI 16) 2016, pp. 265-283