International Journal of Research in Engineering, Science and Management
Volume-1, Issue-12, December-2018
www.ijresm.com | ISSN (Online): 2581-5792

Image Classifications in Pattern Recognition System

M. Parimala
Assistant Professor, Department of Computer Science, Adhiyaman Art and Science College for women, Uthangarai, India

***Abstract*: The copy mode selection, such as the text mode and photo mode, of a digital copy machine can provide suitable process and enhancement for the scanned image. To classify the scanned image without expensive hardware and reduce the running time, in this article, we designed an efficient automatic method for classifying a document image using a probabilistic decision strategy. The proposed algorithm is tailored to inexpensive hardware and significantly reduces both the running time and memory requirements compared to the existing algorithms, while substantially improving the classification accuracy. In addition, we incorporate a new classification module to help avoid moiré patterns by identifying periodic halftone noise.**

***Keywords*: Pattern Recognition**

1. Introduction

A digital copier is a very common piece of home or office equipment. Users typically just push the copy button to make a copy. Most of them are not aware of the fact that copy machines usually have various copy modes associated with different rendering techniques. For example, while the text mode would enhance the edge detail, the photo mode would improve the appearance of very pale colors and smooth the scanned document for noise reduction. Even if the user is aware of different copy modes, it is still cumbersome to select the appropriate copy mode page by page for multi-page documents. Hence, it is essential to develop an automatic page classifier.

The low-complexity method proposed in this paper enables automatic tagging of document images in a low-end copier or all-in-one, by classifying an input original into all possible combinations of mono/color, text/mix/picture/photo, and periodic/stochastic. Note that classifying a document as a photo automatically implies stochastic halftone, hence there is no color-photo-periodic or mono-photo-periodic class. Mono mode is a configuration optimized for monochrome originals while color mode is optimized for color originals. Text mode is optimized for text, line arts, simple graphics, handwritten text, and faxes; picture mode is for high dynamic range halftones originals; photo mode is for continuous tone natural scenes; mix mode is for originals containing both text and picture content; periodic mode is for periodic halftone printed documents, and stochastic mode is for documents printed by other methods.

Misclassifying an original from one class as any of the other classes is an error; however, not all misclassification errors are equally costly. We define two cases of misclassification as benign error: Misclassifying mono originals as color, and misclassifying text or picture or photo originals as mix. All the other misclassification cases are considered harmful errors.

There is a substantial amount of literature related both to the problem of overall segmentation and classification of document images, and to the specific classification tasks considered in this paper. The literature [1], [2] is not applicable to our task due to the stringent complexity restrictions imposed by the low-end machines. Moreover, the document classification algorithms of [3]-[7] access the entire image all at once and visit each pixel multiple times—something that is impossible in the low-end machines.

A number of articles [8]-[10] discussed the related training classifiers. The literature [9] presented the training classifiers using multilayer neural networks to reduce the error in a supervised learning situation. Neural Network techniques can build powerful classifiers with regularization, complexity adjustment and model adjusting. The parameters (weights) in neural network significantly influence the training results. The training analysis in [9], [10] normally is a costly and time-consuming process. The article [11] using multiple instance learning (MIL) to reduce the training instances for handwritten and printed documents classifications. From the results, their scheme can achieve the similar detection accuracy as SVM for the two document image classifications. Nevertheless, the training time and testing time of MIL are still higher than support vector machine (SVM).

The scheme [12] utilizes SVM classifiers with Huffman tree architecture to classify massive documents. The SVM multiple classifiers can be constructed based on Huffman tree with the paragraph and local pixel feature of the input document images. Their scheme can distinguish the texture, character and color from the document images. However, the schemes [11], [12] are complexity and infeasible of distinguishing different modes, such as text, picture, photo, mix, and periodic, for the common scanned image. To classify biomedical document images, extends image classification with scale invariant feature transform (SIFT) by adding color features with bags-of-colors (BoC). In the articles designed a document image classification using convolutional neural network (CNN) that shares weights among neurons among a layer. The schemes aim to distinguish

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-12, December-2018
www.ijresm.com | ISSN (Online): 2581-5792

Table 1
The fourteen distinct classes

	Mono				Color			
	Text	Mix	Pic	Photo	Text	Mix	Pic	Photo
p	mono-text-p	mono-mix-p	mono-pic-p	–	color-text-p	color-mix-p	color-pic-p	–
s	mono-text-s	mono-mix-s	mono-pic-s	mono-photo-s	color-text-s	color-mix-s	color-pic-p	color-photo-s

the content of the input document image, such as the ad, email, news and report. The manner can achieve higher accuracy than by utilizing speeded up robust features (SURF). Consequently, to design an efficient copy mode selection for low-end digital copier, the complexity, time consuming and accuracy should be the major concerns.

In our previous work [1], we demonstrated that our low-complexity image classification algorithms perform with 29 to 99 % accuracy on a large dataset, where misclassifications tend toward benign. Our present work improves upon [1] in two important respects:

Developing new feature extraction and classification methods which result in both lower complexity and higher accuracy than the algorithm of [1]. Specifically:

- We propose a novel classification algorithm. We demonstrate that it improves the classification rate by up to 22 % points as compared to the classifier of [1], when both use the same set of low-complexity features developed in Section.
- We develop a set of features all of which, unlike the features in [1], avoid vertical filtering operations (i.e., computations that involve more than one line of data at a time) and result in 23 and 50 % reductions of the running time and memory requirements, respectively.

We incorporate a periodic halftone classification module developed in which can be added both to the classifier of [1] and to the classifier proposed here, in order to help avoid moiré patterns. Experimental studies in and in Section show that our periodic halftone detector has a 97 % correct classification rate.

2. Algorithm overview and hybrid hard/soft-decision algorithm

We work with a specific copy pipeline equipped with different copy modes which are all possible combinations of mono/color, text/mix/picture/photo, and periodic/stochastic. Our goal is to classify the scanned image of the original into fourteen distinct classes. These classes are listed in the first column.1, where p and s indicate periodic and stochastic, respectively. Note that classes mono-photo-p and color-photo-p are absent, since classifying a document as a photo automatically means stochastic halftone.

In [1], we developed an algorithm for classifying a document as combinations of mono/color and text/mix/photo/picture. That algorithm works by sequentially applying four simple classifiers to a document: first, a classifier to distinguish color from neutral documents; second, a classifier to distinguish text from non-text documents; another classifier to distinguish mix documents from photos/pictures; and a fourth classifier to decide between photos, pictures, and the mix class.

Each classifier i uses a feature vector $\vec{x}$ ix→i consisting of one or two simple features extracted from the document image, and makes its decision based on the decision boundaries shown. The decision boundaries, as well as certain parameters of the feature vectors, are estimated from training data. An additional classifier developed in Figure. It can be added to the classifier [1], as shown.

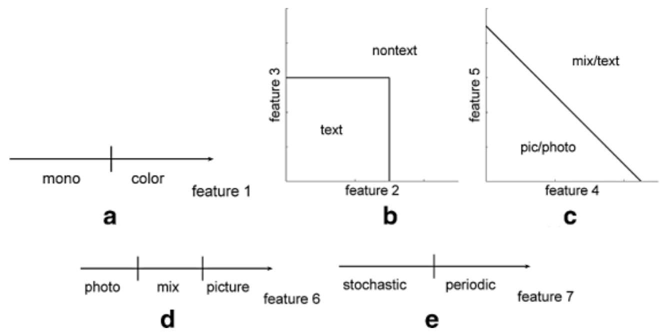

Fig. 1. Decision boundaries for classification nodes

Decision boundaries for classification nodes. (**a**), (**b**), (**c**), (**d**), (**e**) show the decision boundaries for “mono vs. color,” “text vs. nontext,” “text/mix vs. photo/picture,” “mix vs. photo vs. picture,” and “periodic vs. stochastic” classification node, respectively.

A disadvantage of this sequential classification approach is that an incorrect decision made early has no chance of being corrected.

A. *Soft classification algorithm*

As shown, a hard mono-or-color decision is made at the beginning of our new classification strategy. We call the four soft classification nodes shown at the second level nodes 1, 2, 3, and 4, left to right, and let $\vec{x}$ ix→i be the feature vector computed at the *i*-th node. (The computation of feature vectors is described in the next section.)

We let $\vec{X}$ =($\vec{x}$ 1,$\vec{x}$ 2,⋯,$\vec{x}$ n)X→=(x→1,x→2,⋯,x→n) be the overall feature vector obtained from all *n* soft classification nodes: *n*=3 for a and *n*=4 for b. Let c_j, *j*=1,⋯,*M*, be the *M* document classes for the overall classifier, i.e., *M*=8 and *M*=14. Our proposed algorithm estimates the likelihood P ($\vec{X}$ |cj) P (X→|cj) of each class c_j and classifies the document into the class that has the highest estimated likelihood. We assume conditional independence of the feature vectors computed at all nodes, given each class. Hence, each class likelihood factorizes over the *n* classification nodes as follows:

P($\vec{X}$ |cj)=∏iP($\vec{x}$ i|cj).P(X→|cj)=∏iP(x→i|cj).

The class likelihood at each node *i*, P ($\vec{x}$ i|cj) P (x→i|cj), is estimated using a five-bin histogram. The histogram bins are produced for every classifier by using four shifts of the decision boundary in Figure. This is illustrated in Figure for the text-vs.-nontext classifier. Decision boundaries for the soft text-vs-nontext classifier.

Scatter plot of two features used in the text-vs-nontext classification for the color originals in the training suite. *Blue O's* represent text documents, and *red X's* represent nontext documents

Similarly, the outermost bin boundary is chosen to minimize the following number: (number of training non text documents in the outermost bin) - 10 · (number of training text documents in the outermost bin). To obtain the remaining three bins, the distance between the innermost and outermost bin boundaries is then partitioned into three equal parts along each feature axis. To classify a document, we employ a modified maximum likelihood decision rule, constructed so as to bias the decision towards the safe "mix" classification. Given a document to classify, we extract the features, perform the mono-vs-color classification, and estimate the class likelihoods P ($x^{\rightarrow}$ i|cj) P (x→i|cj) at the four soft classification nodes i=1, 2, 3, 4. We then combine these estimates to estimate the overall class likelihoods P ($X^{\rightarrow}$ |cj) P (X→|cj). We classify the document as class j^* if both following conditions hold:

Where, T is a threshold parameter. In our experiments, T=0.85.[2] the first equation corresponds to the standard maximum likelihood classification. The second equation ensures that if there is no clear winner among the different classes, we do not declare a winner.

3. Feature extraction

In this section, we describe all the features used in the four classifier nodes. These nodes use seven features: the mono-vs-color, photo-vs-mix-vs-picture, and periodic-vs-stochastic nodes use one feature each, and the text-vs-nontext and picture/photo-vs-mix/text nodes use two features each.

A. Text vs. nontext classifier

Two features, luminance variability score and histogram flatness score, are utilized to distinguish text documents from nontext documents. We first describe the luminance variability score. We define a *text edge* as five consecutive pixels p_0, p_1, p_2, p_3, and p_4, in horizontal direction, satisfying the following conditions:

- $N(p_1)$, $N(p_2)$, $N(p_3)$ are monotonically increasing or monotonically decreasing,
- $|N(p_1)-N(p_3)|>T_1$,
- $|N(p_0)-N(p_1)|<T_2$ and $|N(p_3)-N(p_4)|<T_2$,

Where, $N(p_i)$ represents the luminance intensity of p_i, and T_1 and T_2 are predefined thresholds. An image block is called a nontext block if there are no text edges in it. To compute the luminance variability score, a test image is partitioned into 8×8 blocks and the mean of each nontext block is calculated.

The second feature, histogram flatness score, is identical to [1], and uses the fact that the histogram for a typical text region has peaks that are more narrow and tall than the peaks in a typical picture or photo histogram. To compute this feature, we partition an image into 8×64 blocks and calculate a 64-bin luminance histogram for each block.

B. Text/mix vs. picture/photo classifier

There are two main differences between text/mix and picture/photo documents: (1) pictures and photos contain no text; (2) pictures and photos contain natural scenes. These two properties are exploited by the two features, the text edge score and the unnaturalness score, that we designed for distinguishing text/mix documents from picture/photo documents.

To describe the text edge score, we first define a *halftone noise triplet* as three consecutive pixels p_0, p_1, and p_2, in horizontal direction, satisfying the following conditions:

- $[N(p_0)-N(p_1)]\times[N(p_1)-N(p_2)]<0$,
- $|N(p_0)-N(p_1)|>T_3$ and $|N(p_1)-N(p_2)|>T_3$,

Where T_3 is a predefined threshold. An image is partitioned into 64×64 blocks. For each block, we count the number of text edges (defined in the previous subsection) and the number of halftone noise triplets.

C. Neutral vs. color classifier

We use the feature for the neutral-vs-color classifier from [1]. We define the colorfulness, $C(p)$, of a pixel p as follows:

C(p)=|I(p)−128|+|Q(p)−128|.C(p)=|I(p)−128|+|Q(p)−128|.

An image is divided into 32×32 blocks.

D. Periodic halftone classifier

We partition the image into 32×32 blocks. For each 32×32 block, we examine every inner pixel, p_{inner}, of the block. We compare the luminance of p_{inner}, $N(p_{inner})$, with luminance values of its four neighbor pixels: $N(p_{left})$, $N(p_{right})$, $N(p_{top})$, and $N(p_{bottom})$. If $N(p_{inner})$ is smaller than any three of the four luminance values from its neighbors, we replace $N(p_{inner})$ with zero. On the other hand, if $N(p_{inner})$ is larger than any three of the four luminance values from its neighbors, we replace $N(p_{inner})$ with 255.

We define region R of the support of $|B_{eh}(u,v)|$ as the union of the following two areas:

- Upper-left: u=(0,1,…,10) and v=(0,1,…,10),
- Upper-right: u=(21,22,…,31) and v=(0,1,…,10).

We let N_R denote the number of points in the region R. Note that the region R excludes the low frequency components region which generally has large coefficients.

4. Experimental results

In terms of memory and time complexity, our approach outperforms [1]. While the text edge and roughness features in [1] require having two strips of data in memory, there is only one strip needed in our algorithm—a 50 % reduction in memory requirements. In addition, since we remove the vertical computations, we also reduce the running time. The average running time per image is approximately 0.268 seconds on an

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-12, December-2018
www.ijresm.com | ISSN (Online): 2581-5792

Table 2
Classification rates for the test data set, using the proposed features

Ground truth	Classification rates, %							
	color-text	**color-mix**	**color-picture**	**color-photo**	**mono-text**	**mono-mix**	**mono-picture**	**mono-photo**
color-text	58/58	42/42	-/-	-/-	-/-	-/-	-/-	-/-
color-mix	-/-	96/98	2/2	2/-	-/-	-/-	-/-	-/-
color-picture	-/-	61/61	39/39	-/-	-/-	-/-	1/1	-/-
color-photo	-/-	36/42	-/-	64/58	-/-	-/-	-/-	-/-
mono-text	13/13	9/9	-/-	-/-	62/56	16/23	-/-	-/-
mono-mix	-/-	9/9	-/-	-/-	1/3	86/86	3/1	-/-
mono-picture	-/-	5/5	6/6	-/-	-/-	40/40	49/49	-/-
mono-photo	-/-	4/4	-/-	2/2	-/-	14/42	-/-	80/58

Intel(R) Core(TM) i7-4770 3.40 GHz desktop for the algorithm. The average running time per image on the same machine for the algorithm of [1] is 0.331 s.

Each entry in the table is “A/B” where A and B are the classification percentages, respectively, for the proposed classifier. Both used with the feature set proposed in the present paper.

We observe that the features proposed in the present paper cause a reduction of the classification accuracies for text, mix, and photo documents. This is due to the fact that our features avoid vertical computations while the ones in [1] do not.

We present the classification results for our proposed hard/soft classification strategy. These are compared to the hard-decision tree classifier. Two experimental results are shown in each entry of the tables using the format “A/B", where A is the classification percentage using the hybrid hard/soft classifier proposed in this paper, B is the classification percentage for the hard-decision tree classifier.

We observe that, at the expense of a very slight reduction in the correct classification rate for color-mix images, our new classification strategy results in significant improvements of the correct classification rates of photo and mono-text documents. Specifically, the hard decision method has 2 % correct classification gain for color-mix, while the proposed hybrid hard/soft method has 6, 6, and 22 % gains for color-photo, mono-text, and mono-photo, respectively.

The two numbers that are more than three percentage points apart are the correct classification rates for mono-picture and mono-photo: the former is 49 % for our algorithm and 30 % for the algorithm in [1], and the latter is 80 % for our algorithm and 66 % for the algorithm in [1].

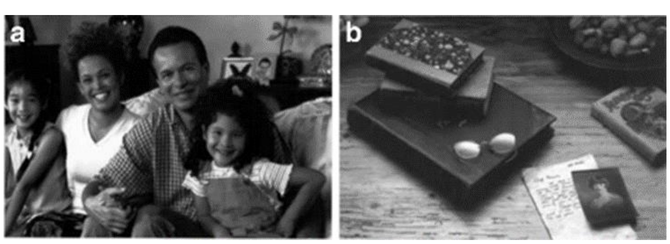

Fig. 2. Mono-photo images

Two mono-photo images that were misclassified by the hard decision method, but correctly classified by our proposed hybrid hard/soft decision method. The hard decision classifier misclassifies them as mix early on in the decision tree and does not even get to compute the roughness feature score which greatly differs between mono-photo and the other mono originals.

Two examples (a, b) that were misclassified by the hard decision classifier, but classified correctly by the hybrid hard/soft decision method

Similarly unaffordable complexity would accompany improvements to our text/mix-vs.-picture/photo classifier. Halftone detection techniques that may be used for separating pictures from photos are discussed in [1]. There is also a vast amount of literature on constructing classifiers [8]-[13]. There exist myriad methods to partition our multidimensional feature space into several classification regions. In designing the overall structure of our algorithm, there were two things we were striving for, besides low complexity and high accuracy:

- Small number of parameters, in order to avoid over fitting.
- Structural simplicity, so that the algorithm is easy to understand and implement.

5. Conclusion

In this paper, we have presented an algorithm to automatically classify documents into a set of categories. This algorithm could be used as a copy mode selector utilized to improve the copy quality and increase the copy rate. As compared to [1], the classification rate is improved by up to 22 % while the running time and memory requirements are saved for 18 and 50 %, respectively.

References

[1] X Dong, K-L Hua, P Majewicz, G McNutt, CA Bouman, JP Allebach, I Pollak, Document page classification algorithms in low-end copy pipeline. J. Electron. Imaging. 17(4), (2008).
[2] W Zhang, X Tang, T Yoshida, Tesc: An approach to text classification using semi-supervised clustering. Knowl.-Based Syst. 75:, 152–160 (2015).
[3] H Cheng, CA Bouman, Document compression using rate-distortion optimized segmentation. J. Electron. Imaging. 10(2), 460–474 (2001).
[4] RL de Queiroz, in Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference On, 1. Compression of compound documents (IEEE, 1999), pp. 209–213.
[5] SV Revankar, Z Fan, in Photonics West 2001-Electronic Imaging. Picture, graphics, and text classification of document image regions (International Society for Optics and Photonics, 2000), pp. 224–228.
[6] SJ Simske, SC Baggs, in Proceedings of the 2004 ACM Symposium on Document Engineering. Digital capture for automated scanner workflows (ACM, 2004), pp. 171–177.

[7] W Wang, I Pollak, T-S Wong, CA Bouman, MP Harper, JM Siskind, Hierarchical stochastic image grammars for classification and segmentation. IEEE Trans. Image Process. 15(10), 3033–3052 (2006).
[8] T Hastie, R Tibshirani, J Friedman, J Franklin, The elements of statistical learning: data mining, inference and prediction. Math Intell. 27(2), 83–85 (2005).
[9] RO Duda, PE Hart, DG Stork, Pattern classification, (2000).
[10] CM Bishop, et al, Pattern Recognition and Machine Learning, vol. 4 (Springer, New York, 2006).
[11] J Kumar, J Pillai, D Doermann, in Document Analysis and Recognition (ICDAR), 2011 International Conference On. Document image classification and labeling using multiple instance learning (IEEE, 2011), pp. 1059–1063.
[12] L Han, J Yu, J Chen, K Zheng, J Luo, Research and application of a method for real estate document image classification. J. Inf. Comput. Sci. 6, 1617–1624 (2012).
[13] M-R Bouguelia, Y Belaid, A Belaïd, in Image Processing (ICIP), 2013 20th IEEE International Conference On. Document image and zone classification through incremental learning (IEEE, 2013), pp. 4230–4234.