

Implementation of Multiple Kernel Support Vector Machine for Automatic Recognition and Classification of Counterfeit Notes

Prof.Sharmishta Desai, Shraddha Kabade, Apurva Bakshi, Apeksha Gunjal, Meghana Yeole

Abstract— With the advance of digital imaging technologies, color scanners and laser printers make it increasingly easier to produce counterfeit banknotes with high resolution. Almost every country in the world face the problem of counterfeit currency notes. Even receiving Fake notes from ATM counters, vending machines and during elections have also been reported at some places. There is a need to design a system that is helpful in recognition of counterfeit notes. In this paper, we propose a system based on multiple-kernel support vector machines for counterfeit banknote recognition. Each banknote is divided into partitions and the luminance histograms of the partitions are taken as the input of the system. Linearly weighted combination is adopted to combine multiple kernels into a combined matrix. Two strategies are adopted to reduce the amount of time and space required by the semi-definite programming (SDP) method. One strategy assumes the non-negativity of the kernel weights, and the other one is to set the sum of the weights to be unity.

Index Terms— Banknote recognition, Support vector machine, Balanced error rate, Multiple-kernel learning, Semi-definite programming, Feature Extraction.

1 INTRODUCTION

Counterfeit banknote recognition system is required in many applications such as automatic selling-goods and vending machines. As a result of the great technological advances in color printing, duplicating and scanning, counterfeiting problems have become more and more serious. Today it is possible for any person to print counterfeit bank notes simply by using a computer and a laser printer at home. The needs for automatic banknote recognition systems encouraged many researchers to develop corresponding robust and reliable techniques.

Several approaches like mask optimization technique using genetic algorithms (GA) and neural networks have been proposed for counterfeit banknote recognition [1]. Support vector machines (SVMs) have been shown to be an effective tool for solving classification problems. The practitioner has to determine the kernel function and the associated kernel hyperparameters in advance. Researchers use trial-and-error to choose proper values for the hyperparameters and this takes lot of efforts. Multiple kernel can solve this problem successfully.

We propose a system based on multiple-kernel support vector machines for counterfeit banknote recognition. For counterfeit banknote recognition, a false positive is usually more harmful than a false negative, since counterfeit banknotes can cause a bigger financial loss if they are not detected. We develop a SVM architecture to favorably reduce false positives. Each banknote is divided into partitions and the luminance histograms of the partitions are taken as the input of the system. Each partition is associated with its own kernels. Linearly weighted combination is adopted to combine multiple kernels into a combined matrix. By applying multiple-kernel

learning, optimal weights with kernel matrices in the combination are obtained through semi-definite programming (SDP) learning. The original SDP problem was formulated on transduction setting where the kernel matrix is created by using the training patterns and the testing patterns. The amount of time and space may grow rapidly as the quantity of data increases. Instead, we consider an induction setting by using only the training patterns to construct the kernel matrix and adopt two strategies to improve the performance of SDP without degrading the accuracy. One strategy assumes the non-negativity of the kernel weights, and the other one is to set the sum of the weights to be unity.

2 SECURITY FEATURES ON INDIAN BANKNOTE

There are various security features in Indian bank note but the most important features mostly used in this paper are given below [4]

1. Watermark

The Mahatma Gandhi Series of banknotes contain the Mahatma Gandhi watermark with a light and shade effect and multidirectional lines in the watermark window.

2. Latent Image

On the obverse side of Rs.1000, Rs.500, Rs.100, Rs.50 and Rs.20 notes, a vertical band on the right side of the Mahatma Gandhi's portrait contains a latent image showing the respective denominational value in numeral. The latent image is visible only when the note is held horizontally at eye level.

3. Fluorescence

Number panels of the notes are printed in fluorescent ink. The notes also have optical fibers. Both can be seen when the notes are exposed to ultra-violet lamp.

4. Microlettering

This feature appears between the vertical band and Mahatma Gandhi portrait. It contains the word 'RBI' in Rs.5 and Rs.10. The notes of Rs.20 and above also contain the denominational value of the notes in microletters. This feature can be seen better under a magnifying glass.

5. Optically Variable Ink

This is a new security feature incorporated in the Rs.1000 and Rs.500 notes with revised colour scheme introduced in November 2000. The numeral 1000 and 500 on the obverse of Rs.1000 and Rs.500 notes respectively is printed in optically variable ink viz., a colour-shifting ink. The colour of the numeral 1000/500 appears green when the note is held flat but would change to blue when the note is held at an angle.

6. See through Register-

The small floral design printed both on the front (hollow) and back (filled up) of the note in the middle of the vertical band next to the Watermark has an accurate back to back registration. The design will appear as one floral design when seen against the light.

7. Serial Numbers

Every banknote has its own serial number, so it is more important to check whether the number is wrong or repeated.

8. Intaglio Printing

This gives a more complex and reliable method, since it is the printing process itself that serves to vouch for the authenticity of the document. The note is subjected to a high-pressure printing process that strengthens and slightly raises the paper's surface structure



Fig 1: Security Features on Indian Banknote

3 Background

3.1. Weighted SVMs

Given a set of training patterns, SVM is a kernel method which finds the maximum margin hyperplane in feature space to separate the training patterns into two groups. To allow for the possibility of outliers in the dataset and to make the method more robust, some patterns need not be strictly and correctly classified by the hyperplane, but the misclassified patterns should be penalized. For this purpose, slack variables ξ_i are introduced to account for the misclassified patterns. The objective function and constraints of the problem can therefore be formulated as:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + c \sum_{i=1} \xi_i \tag{1}$$

$$s.t. y_i ((w, \phi(x_i)) + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, \dots, l.$$

In the following experiments, we use $C = 1$ for standard SVMs and $C = 2l+$ for our proposed SVMs of eqn

$$\min_{w,b} \frac{1}{2} \|w\|^2 + c \left(\sum_{\{i|V_i=+1\}} d_i + \xi_i + \sum_{\{i|V_i=-1\}} d_i - \xi_i \right) \tag{2}$$

$$s.t. y_i ((w, \phi(x_i)) + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, \dots, l.$$

As it is well known, SVMs are trained through the following optimization procedure by using the equations,

$$\text{Polynomial: } K(x, x') = (x, x'+1)^d \tag{3}$$

$$\text{RBF: } K(x, x') = \exp(-\gamma|x - x'|^2) \text{ for } \gamma > 0 \tag{4}$$

where $K(\cdot)$ is called the kernel function. The value of the kernel equals the inner product of two vectors, x_i and x_j , in the feature space $\Phi(x_i)$ and $\Phi(x_j)$; that is $K(x \& x') = \Phi(x_i) \cdot \Phi(x_j)$. In this work the RBF is used. Here γ is width parameter of RBF kernel[3].

3.2. Kernel fusion and weight learning

Early SVM-based methods used a single-kernel function, $K(x \& x') = \Phi(x_i) \cdot \Phi(x_j)$ to calculate the inner product between two images in the feature space F . In practice, the kernel function is characterized by a kernel matrix computed from the training patterns. A kernel matrix is a square matrix where each matrix entry measures the similarity between $F(x_i)$ and $F(x_j)$. If a dataset has varying local distributions, using a single

kernel may lose some information to cope with this varying distribution. Most researchers use the trial-and-error heuristic to choose the best hyperparameters, which obviously takes a lot of efforts. Kernel fusion can help to solve this problem. A simple direct sum fusion can be defined as $K(x \& x') = (\Phi(x), \Phi(x'))$, where Φ is a new feature mapping. This feature mapping can handle issues of varying pattern distributions by using multiple-kernel functions. The kernel matrix can be easily written as $K = K_1 + \dots + K_M$ in this case, with K_i obtained from Φ_i .

This simple fusion can be generalized to a weighted combination of kernel matrices as follows:

$$K = \sum_{s=1}^M \mu_s K_s \quad (5)$$

where M is the total number of kernel matrices and μ_s is the weight of the s th kernel matrix.

3.3. Proposed work

The proposed system will work on two images, one is original image of the paper currency and other is the image on which verification is to be performed.

The proposed Algorithm is as followed:

1. Image of currency will be captured by scanner or digital camera.
2. Segmentation of given image will be carried on. To cope with varying histogram distributions in different areas of a banknote, each banknote is divided into $m \times n$ non-overlapping partition.
3. Feature extraction is performed by transforming the input data into set of features. Each partition of given image is represented by the luminance histogram of that partition.
4. With the training patterns, kernel matrix is formed.
5. In the testing phase, our system is applied to determining whether a given banknote is genuine or fake.

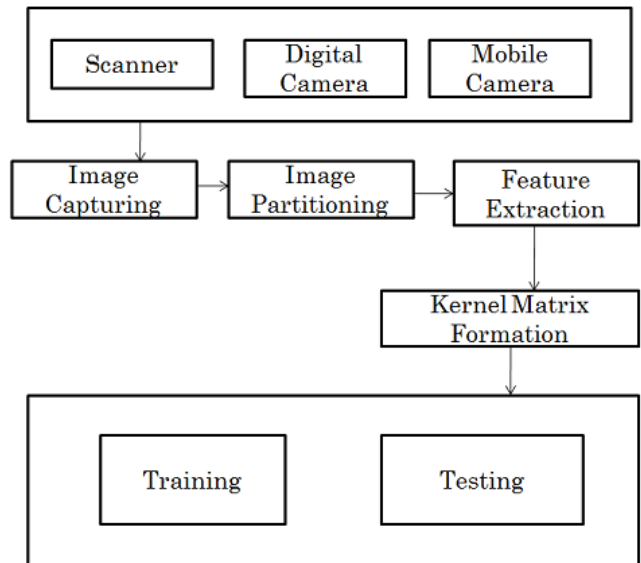


Fig 2: System Architecture

4 Proposed SVM architectures

4.1. Single-kernel weighted SVM

The single-kernel SVM designed for our purpose can be formulated as given in eq (2) where $y_i = +1$ denotes that x_i is a genuine banknote while $y_i = -1$ denotes that x_i is a counterfeit banknote. Our proposed SVM can minimize the balanced error rate.

4.2. Multiple-kernel weighted SVM

The amount of time and space requirements grows rapidly as the quantity of data increases, since K is obtained from the training patterns and the testing patterns. One needs to check whether the combined matrix is positive semi-definite in each iteration. This introduces huge search space for finding optimal solutions. To reduce the computational complexity, we consider an induction setting by using only the training patterns to construct the kernel matrix. In addition, we adopt two strategies to help narrow down the search space for kernel weights. The first strategy is to assume the non-negativity of kernel weights and the second strategy is to set the sum of weights equal to 1.

5 Experiment Results

5.1. Experiment I

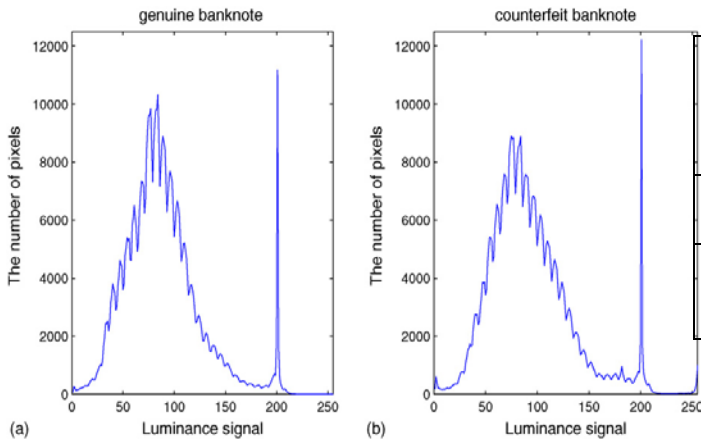


Fig 3: Genuine currency vs Counterfeit currency

Note that in the table, Sd-Sk indicates that it is standard SVM with single kernel. Op-Sk stands for our proposed SVM with single kernel. FPR stands for false positive rate. FNR stands for false negative rate. ACC stands for accuracy. From Table 1 we can see that our proposed SVMs performs better than standard SVMs. At $\gamma = 0.165$, Sd-Sk-Np achieves the highest accuracy rate of 83.871%, while at $\gamma = 0.2$, Op-Sk-Np achieves the highest accuracy rate of 87.097%.

TABLE 1

Comparison between different SVM architectures (maximum values indicated in boldface).

kernel	Sd-Sk ACC(%) / FPR(%) / FNR(%)	Op-Sk ACC(%) / FPR(%) / FNR(%)
$\gamma=0.01$	67.742/90.909/0.000	70.968/45.455/20.000
$\gamma=0.1$	74.194/54.546/10.000	77.419/36.364/15.000
$\gamma=0.05$	74.194/54.546/10.000	74.194/45.455/15.000
$\gamma=0.5$	80.645 /45.455/5.000	80.645 /45.455/5.000

5.2. Experiment II

In this experiment, we investigate the effect of using multiple kernels. No partitioning is done. The results are shown in Table 2 for two sets of kernels, 3-kernels and 5-kernels. In the table, Op-Mk-Tr stands for our proposed SVM with multiple kernels and the kernel weights are learned by SDP with transduction, while Op-Mk-Id learns by SDP with induction. From Tables 1 and 2, we can see that using multiple kernels helps improve the accuracy rate, from 80.645% to 83.871%. However, the accuracy rate does not change noticeably with different sets of kernels. Note that our proposed approach improves the time efficiency of the original SDP method without degrading the performance.

TABLE 2

Results obtained from combining multiple kernels.

RBF kernel	Op-Mk-Tr ACC (%) / FPR (%) / FNR (%) / time (s)	Op-Mk-Id ACC (%) / FPR (%) / FNR (%) / time (s)
$\gamma = [0.050.55]^T$	83.871 /36.364/5.000/15.074	83.871 /36.364/ 5.000/10.646
$\gamma = [0.010.050.50.1]^T$	80.645/45.455/5.000/ 15.960	83.871 /36.364/ 5.000/13.046

5.3. Experiment III

In this experiment, we investigate the effect of applying the strategy of partitioning. Tables 3 and 4 present the results for two sets of kernels, $\gamma = [0.050.55]^T$ and $\gamma = [0.010.050.10.51]^T$, respectively. In these tables, Op-Mk-Tr-Pa and Op-Mk-Id-Pa stand for Op-Mk-Tr and Op-Mk-Id, respectively, with partitioning. Clearly, different ways of partitioning may result in different performances. For example, partition 2×2 gets 100% in accuracy for Op-Mk-Id-Pa with 3-kernels, while partition 2×4 gets a poor performance, only 80.645% in accuracy. Also, a partition may behave differently with different sets of kernels. For example, partition 4×2 gets 96.774% in accuracy for Op-Mk-Id-Pa with 3-kernels, while it gets 100% in accuracy for Op-Mk-Id-Pa with 5-kernels. The variation in performance indicates the variation in histogram distribution due to different ways of partitioning. Again, Tables 3 and 4 show that our proposed SDP method improves the time efficiency of the original SDP method.

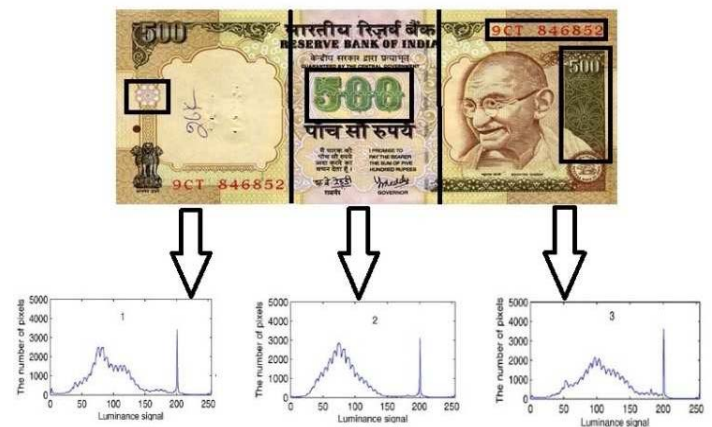


Fig 3: Partition of banknote with histogram

TABLE 3

Results obtained from partitioning for $\gamma = [0.05 \ 0.55]^T$.

RBF kernel	$\gamma = [0.05 \ 0.55]^T$	
Methods Partitions	Op-Mk-Tr-Pa ACC (%)/FPR (%) /FNR (%) /time (s)	Op-Mk-Id-Pa ACC (%)/FPR (%) /FNR (%) /time (s)
2 x 2	96.774 /9.091/0.000/17.271	100.000 /0.000/0.000/11.025
2 x 4	80.645/54.546/0.000/2 1.683	80.645/54.546/0.000/12 .270
4 x 2	96.774 /9.091/0.000/22.974	96.774/9.091/0.000/12. 139
4 x 8	90.323/27.273/0.000/3 5.968	93.548/18.182/0.000/14 .138
8 x 8	77.419/63.636/0.000/9 3.182	96.774/9.091/0.000/31. 559
8 x 16	87.097/36.364/0.000/2 18.608	100.000 /0.000/0.000/71.629

TABLE 4

Results obtained from partitioning for $\gamma = [0.01 \ 0.05 \ 0.10 \ 0.51]^T$.

RBF kernel	$\gamma = [0.05 \ 0.55]^T$	
Methods Partitions	Op-Mk-Tr-Pa ACC (%)/FPR (%) /FNR (%) /time (s)	Op-Mk-Id-Pa ACC (%)/FPR (%) /FNR (%) /time (s)
2 x 2	93.548/18.182/0.000/18 .806	100.000 /0.000/0.000/11.815
2 x 4	87.097/36.364/0.000/24 .418	80.645/54.546/0.000/13.4 03
4 x 2	100.000 /0.000/0.000/26.111	100.000 /0.000/0.000/13.656
4 x 8	93.548/18.182/0.000/39 .462	93.548/18.182/0.000/16.8 78
8 x 8	80.645/54.546/0.000/17 7.823	90.323/27.273/0.000/53.3 98
8 x 16	80.645/54.546/0.000/17 7.823	93.548/18.182/0.000/149. 829

6 Advantages and Disadvantages

6.1. Advantages of SVM

1. It has a regularization parameter, which makes the user think about avoiding over-fitting.

2. Produce very accurate classifiers.
3. It is defined by a convex optimization problem.

6.2. Limitations of SVM

1. The performance of SVMs largely depends on the choice of kernels.
2. SVM is a binary classifier.
3. Computationally expensive.
4. Long training period and recognition rate.

7 Conclusion

In this paper a method based on multiple-kernel support vector machines for counterfeit banknote recognition is explored. Our main aim behind this paper was to present the system based on recognition of fake currency banknotes to avoid frauds. We have considered Indian banknotes as references. Each banknote is divided into partitions and then histograms of partitions are taken as input. To combine multiple kernels into a combined matrix, linearly weighted combination is adopted. By applying multiple-kernel learning, optimal weights with kernel matrices in the combination are obtained through semi-definite programming (SDP) learning. A method is proposed which can narrow down the search space for searching optimal parameter settings. Only training patterns are used to construct kernel matrices to reduce the computational complexity.

Our system will be able to distinguish between genuine and forged banknotes without any modification even if more counterfeit security features are added to the banknotes. Our system can be used in banks, shops and in many other places.

8 References

- [1] F. Takeda, T. Nishikage, S. Omatu, "Banknote recognition by means of optimized masks, neural networks and genetic algorithms", Engineering Applications of Artificial Intelligence Vol: 12 no:2 pp: 175-184,(1999).
- [2] Sanjana, Manoj Diwakar, Anand Sharma, "An Automated Recognition of Fake or Destroyed Indian Currency Notes in Machine Vision", (2012).
- [3] Mohd Mudassir Mohd Ismail, Mr. Ashutosh. B. Makone, "Survey of An Automated Recognition of Fake or Destroyed Indian Currency Notes", (2014).
- [4] Sebastiano Battiato, Giovanni Maria Farinella, Arcangelo Bruna and Giuseppe Claudio Guarnera, "Counterfeit Detection and Value Recognition of Euro Banknotes", (2013).
- [5] Sharmishta Desai, Dr.S.T.Patil, "Automatic Detection and Classification of Objects with Optimized Search Space", (2014).
- [6] Chi-Yuan Yeh, Wen-Pin Su, Shie-Jue Lee, "Employing multiple-kernel support vector machines for counterfeit banknote recognition", pp:1439-1447,(2011).
- [7] H. Drucker, D. Wu, V. Vapnik, "Support vector machines for spam categorization," IEEE Transactions on Neural Networks vol:10 no:5 pp: 1048-1054,(1999).