

Recognition System for Euro and Mexican Banknotes Based on Deep Learning with Real Scene Images

Deysy Galeana-Pérez, Eduardo Bayro-Corrochano

CINVESTAV, Department Electrical Engineering and Computer Science, Guadalajara, Mexico

{dgaleana, edb}@gdl.cinvestav.mx

Abstract. This article presents a robust and efficient system for euro and Mexican banknote recognition. A high banknote recognition and classification rate was achieved using neural networks and deep learning with real scene images taken with both sunlight and artificial light. Without extracting characteristics by hand, the convolutional neural networks was fed with raw images. Analysis and experiments were carried out on banknotes based on key features, such as; watermarks, portraits on the bills, bill value written in words and numbers, and the complete banknotes. It was concluded that both the color information and some regions of the banknotes, as well as the banknote denomination written in words and numbers and the complete banknote, is the appropriate information to achieve a high rate of banknote classification and recognition. The experimental results show that the proposed approach is promising with quite remarkable results; it performs an efficient and robust classification using real scene images taken with both sunlight and artificial light and is invariant to banknote rotation and translation. A high recognition rate was achieved for Mexican banknotes and for euros. At present, the results contained herein are an improvement over those reported in the state of the art.

Keywords. Banknote recognition, convolutional neural networks, deep learning.

1 Introduction

Banknote recognition and classification is a challenge in the area of computer vision and automatic learning, as there is no methodical process for extracting the most relevant features in banknotes. Research into better algorithms

for extracting these characteristics has increased, as a proper selection of characteristics makes the classification algorithms successful. Better methods are currently being investigated to recognize and classify objects of different shapes, textures, and colors contained in images. The main objective of automatic learning is to develop algorithms that help the computer to learn under certain data inputs, in our case, images on banknotes, and the algorithm would focus its learning on the most relevant parts of the banknotes.

The problem of banknote recognition is that all bills are rectangular. For a good rating and recognition, it is better to focus on the color and texture information by taking parts of the bill to analyze. For artificial vision, the banknotes are considered objects that share the same shape (rectangular in all countries) and can be differentiated by the color and texture of each for their recognition and classification.

Research into banknotes has increased due to the demand in the financial sector for Automated Teller Machines (ATMs) and currency counting machines, in the commercial sector for automated vending machines and automated ticket vending machines, and in the social sector for assisted currency recognition for blind or visually impaired people. Robust systems are required in all these fields, leading to this proposal for better algorithms and methods for banknote recognition and classification.

Methods and algorithms have been proposed in the field of banknote recognition to detect

counterfeit banknotes, as presented in [30], to classify new and old banknotes [28, 13], identification of banknote denomination to help the visual impaired [22], classification of banknotes from different countries [17, 18, 31], and algorithms proposed to detect defects in banknotes [5].

The methods used to extract the banknote characteristics are quaternion wavelet transform (QWT) and generalized Gaussian density, RGB color space and Local Binary patterns; these features are usually used as input to a neural network, which is trained to recognize and classify the banknotes.

In the approach here proposed, experiments were first conducted to determine the best banknote regions for network training. Image preprocessing and manual feature extraction are not necessary, as deep learning circumvents manual feature extraction; the network learns the best features to represent the data provided in each layer. The Deep Learning Network extracts feature maps from each layer; these maps contain information about the edges, corners, shapes, and colors of objects found in the image and network learns from these feature maps.

The objective of this article is to show the benefits of deep learning applied to banknote recognition and classification by providing the network with the best regions for learning. A robust and efficient system which is used in this task a great challenge in artificial vision, due to the problems of illumination, occlusion, background disorder, rotation, translation, variations in point of view, deformation, and object depth.

The rest of this paper is organized as follows: Section 2 presents the state of art of banknote recognition. Section 3 presents a discussion about convolutional neural networks and deep learning. Section 4 introduces the banknote recognition system. Section 5 discusses neural networks based on perception systems for banknote recognition. Section 6 presents the experimental analysis and Section 7 offers conclusions and discusses future work.

2 Related Work

Several techniques or algorithms for banknote recognition have been used in recent years for recognizing and classifying banknotes from several countries, such as Chinese renminbi, Japanese yen, Italian lira, euros, American dollars, Colombian pesos, and Mexican pesos, among others. This technology is envisaged to help financial institutions recognize banknotes for accounting, for ATMs, and in the identification of new and old or damaged banknotes, as well as to help identify counterfeit banknotes in business and in social transactions, e.g., to assist the visually impaired. In [28], an algorithm is presented for classifying new and old banknotes based on statistical parameters of the characteristic points based on the cooccurrence matrix of the gray levels with favorable results.

In [7], the RGB space and Local Binary patterns are used to extract color and texture descriptors, respectively. These descriptors are the inputs to a Linear Vector Quantization (LVQ) network and G statistic; this network classifies the banknotes. This method is not invariant to illumination though. Feature extraction was performed manually, and it was demonstrated that by combining the characteristics of color and texture, the classification significantly improved over those using only color or texture descriptors. In [22], a method is presented for identifying Mexican banknotes based on geometric pattern identification without any training or classifier, obviating the need to build a training database. For the extraction of geometric patterns, the Fast Radial Symmetric transform (FRS) is used for detecting the semicircular elements of the numbers in the banknote denominations.

Other researchers have employed QWT for the recognition of banknote denominations. The authors [17] perform a preprocessing to correct the tilt and brightness of an image with a black background; these images are fed to the neural network input. The features were extracted using QWT and the generalized Gaussian density describes the characteristics of the transformed coefficients; this vector is in turn the input to the neural network BackPropagation.

The performance of the proposed method returns a rate of 99.90%, 99.65%, and 99.50%, in dollars, RMBs, and euros; respectively, in the training database. However, the approach needs to capture the image in a fairly controlled environment.

A new algorithm is proposed in [5] to detect cracks and scratches on banknote images by means of the QWT. Experiments were carried out on a database for Chinese banknotes, dollars, and euros, showing favorable results and a high recognition rate.

Furthermore, [18] proposes a banknote recognition method based on discriminative region selection by using a method of masking on the banknote image capture by a visible light sensor. Masks are randomly initialized from a selection priority map derived from a similarity map that represents the region classification capacity in the undersampled banknote image. The genetic algorithm is used for mask optimization to obtain a better recognition region selection. For the database, Hong Kong dollars, Kazakhstani tenges, Colombian pesos, and American dollars are used, achieving a rate of 99.386%, 100%, 99.726%, and 99.990%, respectively.

Banknote size has also been used in [31] to propose an algorithm that takes information on banknote size as the most important characteristic, the authors considered a preprocessing to the banknote images using a two-dimensional Weiner filter to reduce the noise factors generated by acquiring the images via a contact image sensor (CIS). After the scan, a bias is applied to correct the rotation and extraction of the ROI (region of interest) to separate the areas from the banknotes. The experimental results show a 100% and 99.8% classification accuracy for clean and stained banknotes, respectively.

Other work has focused on detecting the degree of damage to banknotes by classifying old and new banknotes. In [13], the Hough transform is used to align the banknotes and, extract relevant characteristics, such as the denomination or watermarks, and a Support Vector Machines (SVMs) is then used for the classification of old and new banknotes.

Others have focused on detecting serial numbers on banknotes, carrying out a preprocessing

on each image to locate and segment the area with the serial number. In [6], an algorithm is proposed to identify serial numbers on bills based on SVM. These experiments were carried out on 3 databases: American dollars, euros, and Chinese RMBs. There were eight thousand samples for six denominations and four facial positions on the banknotes (up, down, reversed up, reversed down), with 24 classes for the dollars.

Euros have 28 classes, 7 denominations, and 4 different facial positions, and RMB have 5 different denominations of Yuan and 4 different sides to each banknote, for a total of 20 different classes. Recognition rates of 99.34%, 99.18%, and 99.64% were obtained, respectively. In [27], an optimized masking technique is proposed that uses genetic algorithms for a serial number recognition system by means of neural networks.

The system was tested on Italian lira banknotes with an effectiveness of 97%. There are several feature extraction techniques for providing these characteristics to a classifier. Our deep learning approach uses the same algorithm to extract representative characteristics from each object class, providing the algorithm with entries and, images with raw pixels, without the need for prior pixel processing. Deep learning applied to neural networks will be discussed in the following section.

3 Deep Learning

The objective of artificial intelligence systems is to help a computer learn by designing a set of algorithms to extract their own knowledge, i.e., to extract representative patterns of raw data, an ability known as machine learning. Deep learning is a branch of machine learning that enables a computer to learn from experience, constructing complex concepts from simpler concepts [23, 9]. For example, the concept of banknotes is first built by simpler concepts, such as corners and contours, which in turn represent the edges of the banknotes. These edges form the figures of the banknotes belonging to the reliefs, watermarks, portraits, denominations in number and words, etc. This process is carried out through several layers of depth implemented in the machine learning algorithms.

Each layer learns the best characteristics to represent the given data. These features are extracted directly from images, text, or raw sound, eliminating the need for manual extraction. One of the most popular deep learning algorithms is convolutional neural networks (CNN) [16, 15]. In 2012 Alex Krizhevsky [14] created a large, deep convolutional neural network to classify high-resolution images in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC-2010). This network won the contest, giving it a huge boost, and several companies started using it. Facebook uses CNN for its automatic tagging algorithm, Amazon for its product recommendations, and Instagram for its search infrastructure [4], but its most popular use is in the field of image processing [3].

CNNs have been specialized to work with images, using filters or kernels that perform the convolution process, making it possible to detect and highlight edge information useful for segmenting the objects contained in an image. The human visual system can easily and effortlessly identify its environment and nearby objects. Humans can immediately characterize a scene and assign a label to each object in their environment without thinking about it. When a computer sees an image, i.e., it takes an image as input, what it sees is an array of pixel values from 0 to 255 that correspond to pixel intensity.

These are the only inputs available for the computer, and from these, the computer will be able to differentiate between all the arrays of pixel values (images) given, looking for low-level features, such as edges and curves, and then building more complex concepts through a series of convolutional layers [9]. The computer then discovers how the unique features differentiate one banknote from another and assigns a probability of it belonging to a certain class. The layer does not need to learn the whole concept at once; each layer gathers different information by making a chain of features that build that knowledge [23], (See Fig. 1).

It is worth noting that a convolutional neural network has not only convolutional layers, but also pooling, dropout, batch normalization, fully-connected, relu, tanh, sigmoid, softmax,

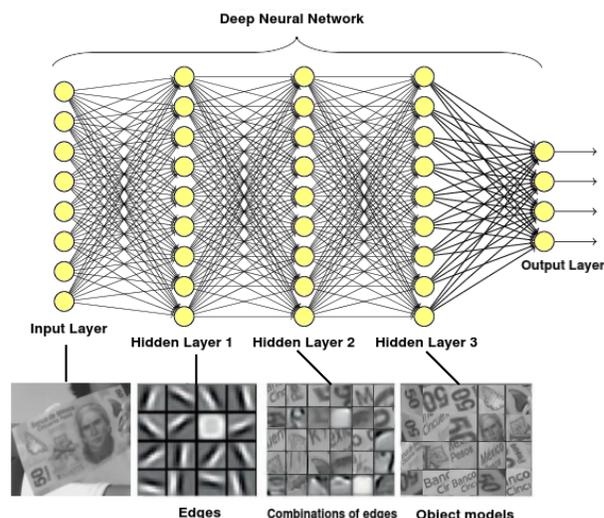


Fig. 1. Each layer learns a concept with the output from the previous layer, forming complex model objects in deeper layers

cross entropy, SVM, and Euclidean layers as well. Overfitting is a serious problem in networks [29] and prevents a network from learning. Dropout is a technique for addressing this problem [25], the key idea is to randomly drop units (along with their connections) from the neural network during training, preventing units from excessive co-adapting and significantly reducing overfitting and greatly improving upon other regularization methods. Dropout can improve the performance of neural networks in supervised learning tasks in vision, speech recognition, document classification, and computational biology [2, 1].

CNN has been used primarily for object detection, location, and classification in images [14, 32, 24, 26, 11] and CNN based on regions [8, 21] uses a combination of CNN and RNN to generate descriptions in natural language in the image regions.

This work presents a system to detect and recognize banknotes from different denominations and from different countries; Mexicans and euros use convolutional neural networks with deep training. The network training was carried out using images taken with both natural and artificial lighting, with variations in the lighting, rotation, and

translation of the banknotes, as well as at different distances and depths. It is worth noting that unlike in other works, images in this database were grouped into a single class (up, down, reversed up, reversed down), so that the system could detect and recognize the banknote denomination regardless of its position as a single banknote class.

4 System Perception Banknotes

In this system, banknote perception is based on neural networks. A novel method is described below that allows improving the identification of banknotes from different countries to a degree equal to or greater than the results previously presented in the state of the art. This system is invariant to banknote position (banknotes have four faces, which are up, down, reversed up, reversed down). It is not necessary to train each of these positions separately; they are taken as a single class.

During the state of the art analysis, it was concluded that there are key factors that improve the banknote identification process, such as the amount of light falling on the object, the orientation of the banknotes, and the quality of the camera. However, in addition to the above mentioned variables, others are proposed here, such as banknote denomination, color, watermarks, and portrait shape, thus increasing the number of training images. Of the different techniques used, neural networks present the best results. Neural networks with deep learning or deep convolutional neuronal networks are applied, as the state of the art has shown these to work best with images.

The architecture of the object detection network used is based on [20]. It takes an image as input, passes it through a CNN, and receives a vector of bounding boxes and class predictions in the output. The input image is divided into an $S \times S$ grid of cells. For each object present on the image, one grid cell is responsible for its prediction: the cell into which the center of the object falls.

Each grid cell predicts B bounding boxes and C class probabilities. The bounding box prediction has 5 components: (x, y, w, h, confidence).

The (x, y) coordinates represent the center of the box relative to the grid cell location [20]. These coordinates are normalized to fall between 0 and 1. The (w, h) box dimensions are also normalized to [0, 1], relative to the image size, so there are $S \times S \times B * 5$ outputs in total related to bounding box predictions. It is also necessary to predict the class probabilities, $Pr(Class(i)|Object)$. This probability is conditioned on the grid cell containing one object, i.e., if no object is present in the grid cell, then the loss function will not penalize it for an incorrect class prediction.

The network only predicts one set of class probabilities per cell, regardless of the number of boxes B, resulting in $S \times S \times C$ class probabilities in total [20]. The network has 24 convolutional layers and 2 fully-connected layers. The final layer uses a linear activation function and the remaining layers use a leaky RELU activation function (equation 1) with parameter α equal to 0.1 [10]. This paper is based on a fast version, with fewer convolutional layers and some modified parameters:

$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha x & \text{otherwise.} \end{cases} \quad (1)$$

The loss function is composed of several parts [20]:

$$\begin{aligned} & \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 + \\ & + \lambda_{coord} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] + \\ & + \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} (C_i - \hat{C}_i)^2 + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{noobj} (C_i - \hat{C}_i)^2 + \\ & + \sum_{i=0}^{S^2} 1_i^{obj} \sum_{c \in classes} (p_i(c) - \hat{p}_i(c))^2. \end{aligned} \quad (2)$$

In [20], the authors explain this equation as follows: the first part of equation 2 computes the loss related to the predicted bounding box position (x,y), i.e., a sum over each bounding box predictor ($j = 0..B$) of each grid cell ($i = 0..S^2$). And 1_{ij}^{obj} is 1 if an object is present in grid cell i and the j-th bounding box predictor is responsible for that prediction.

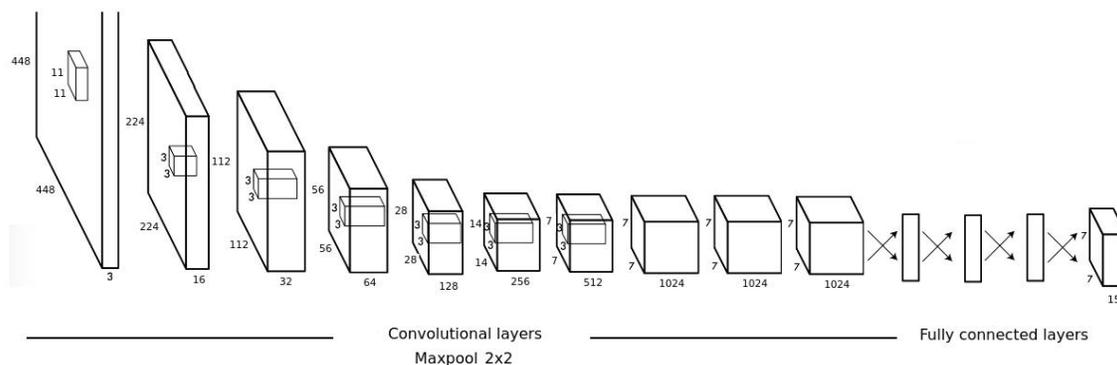


Fig. 2. Architecture of the network used

(x, y) are the predicted bounding box positions and (\hat{x}, \hat{y}) are the current positions from the training data.

The second part corresponds to the error metric which should reflect that small deviations in large boxes are less important than those in small boxes. To partially address this, the square root of the bounding box width and height is predicted instead of the width and height directly.

The third part corresponds to the loss associated with the confidence score for each bounding box predictor is computed. C is the confidence score and \hat{C} is the Intersection over Union of the predicted bounding box with the ground truth. 1^{obj} is equal to one when there is an object in the cell (increasing the confidence to the Intersection over Union), and 0 otherwise (decreasing the confidence to 0). 1^{noobj} is the opposite.

The last part of the loss function is the object classification loss. The term 1^{noobj} is used to penalize a classification error when no object is present in the cell. The cost function presented in [20] is necessary to carry out the training. The architecture of the network used for training this system was inspired by [20] and is shown in Fig. 2. The CNN architecture here presented consists of 9 convolutional layers with 3×3 kernels.

The number of kernels increases twice in each convolutional layer; the first layer has 16 kernels and the final 3 layers have 1,024 kernels. After each convolutional layer is a maxpool operation of 2×2 over the remaining size of the image, and with a leaky ReLU operation function in each

convolutional layer, and 3 fully-connected layers with a linear operation function and a detection layer at the end.

The network was trained for 65 epochs using the training data sets, with a batch size of 64, a momentum of 0.9, and a decay of 0.0005. The rate dropout was 0.5 before the final layer was fully connected to prevent co-adaptation on the training data [12]. That is, each hidden unit is randomly omitted from the network with a probability of 0.5, the parameters $\lambda_{coord} = 5$ and $\lambda_{noobj} = 0.5$, and the input resolution of the network from 448×448 . The image was divided into a 11×11 grid and for each grid cell, 2 bounding boxes and $C = 5$ were predicted for Mexican banknotes and $C = 7$ for euro banknotes.

4.1 Training Data

A prior analysis was performed to select parts to segment the banknotes, to learn about the key features, which succeeded in separating each group or banknote class effectively and efficiently. This analysis was performed using 300 Mexican banknote images by denomination. Samples are shown in Fig. 3. This analysis was performed with 3 proposed regions from various parts of the banknote.

This method of region selection was then extended to euros. Fig. 4 shows samples from this image database and Fig. 5 shows the regions selected for the euro banknotes to train



Fig. 3. Denomination samples for the Mexican pesos database



Fig. 4. Denomination samples for the euro database

the network. The first proposed the complete banknote, as shown in Fig. 5a.

The second proposed all objects on the banknote, i.e., the faces in the portraits appearing on banknotes, the amount in numbers and words, the name of the banknotes, and symbols contained on the banknotes, (see Fig. 5b).

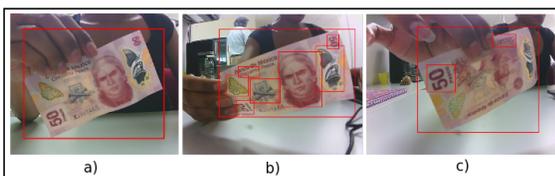


Fig. 5. Samples of segmented images. a) Segmenting only the banknotes, b) segmented symbols, the faces in the portraits, and the denomination of the banknotes in numbers and words, c) Segmented banknotes and banknote denomination written in words and numbers

The third proposal was to leave the complete banknote and banknote denomination (see Fig. 5c). The latter proposal showed better results than the former.

This analysis was carried out with color and grayscale images to verify which image database provided better classification results and to ascertain whether color information was helpful in classification. Once it was determined which banknote regions had better performance for the classification of Mexican banknotes, those regions were then selected for euros (see Fig. 6). Prior to this method for proposing regions, tests were carried out to binarize the images and subtract the average from each of the images used in training, for homogeneous images with respect to brightness. No acceptable results were obtained with this image preprocessing.

4.2 Preparation of the Proposed Regions

As noted above, the method for selecting regions with acceptable results was to carry out experiments with the same set of images in, grayscale and color and with each of the proposed regions or items on the banknotes and then validate the results.



Fig. 6. Regions selected from the images of euros

Images without preprocessing are provided to the convolutional neural network along with the text files and the information on the coordinates of the objects found on the image.

5 Perception System Based on Deep Neural Networks

Deep learning is a branch of automatic learning based on a set of algorithms that learn to represent the data, with the most popular algorithms being convolutional neural networks, deep belief networks, deep auto encoders, and recurrent

neural networks with Long Short-Term Memory units, or LSTMs [23]. The latest trend is to use neural networks that mimic how the human brain learns to identify and label objects presented to it, i.e., by extracting the shapes, edges, or textures of the panorama presented to it.

In this article, convolutional neural networks are used for the classification of banknotes, labeling the boxes into which the input image is divided into in the network, in this case 11×11 . Each cell has the probability of being one of the 5 or 7 classes of Mexican peso or euro banknotes, respectively, and the network also performs a segmentation of the object or banknote, enclosing it with a bounding box in the image.

An analysis was carried out to select the network architecture to be used. Convolutional neural networks have been used for image classification [3, 20] and text classification [33]. To select the network used in this work, several architectures were tested and the model from [20] was selected by reducing convolutional and fully-connected layers, as well as the number of filters used in each convolutional layer. The same leaky activation function was used. Some architectures used did not function, as the gradient was undetermined, and so were discarded. Fig. 7 shows the model of the network architecture used for the recognition of Mexican peso and euro banknotes.

The network used to classify banknote denominations is composed of 9 convolutional layers followed by 2×2 Maxpool layers, 3 fully-connected layers and an exit or detection layer. This architecture is based on [20] and trained in the darknet framework [19]. The memory used for each image is approximately 34.6 MB and the parameters for each image for forward propagation alone is approximately 46.5 million.

The output of this model is a $7 \times 7 \times 15$ tensor for the Mexican banknote network and a $7 \times 7 \times 17$ tensor for the euro bill network. This tensor can be considered a 7×7 grid representing the input image, where each tensor cell will have the 2 box definition. Each box contains the parameters of x , y , width, height, and the confidentiality of objects belonging to the classes.

The network was trained on a NVIDIA Geforce GT 740 SC graphics card (2GB VRAM), mounted

in a desktop computer with a i7 3.50 GH core processor with 8GB in RAM for approximately 5 days. The following section outlines the experiments carried out once the network finished training.

6 Experimental Analysis

Several experiments were carried out to select the set of relevant regions on the banknotes, allowing a successful training with a greater generalization capacity. A total of 1,500 images were used to train in these 3 experiments, with 300 per class. Table 1 shows the results of the network trained with entire banknote. (See Fig. 5a).

Table 1. Taking the entire banknote as a proposed region

Mexican banknotes	Controlled		Semi-controlled		Not Controlled	
	gray	Color	gray	Color	gray	Color
20 pesos	50%	90%	6%	8%	0%	98.33%
50 pesos	16%	60%	8%	18%	7.5%	4.16%
100 pesos	90%	94%	98%	100%	81.66%	100%
200 pesos	74%	86%	18%	100%	8.33%	100%
500 pesos	0%	80%	64%	100%	6.66%	0%
Efficiency	46%	82%	38.8%	62.8%	20.83%	60.50%
Glb eff. gray	35.21%					
Glb eff. color	68.43%					

In this case, the global efficiency with grayscale images is 35.21% and the color is 68.43%. The network has a low generalization capacity for new samples here. There is considerable confusion between classes and more still in grayscale images, where color is discriminated against. Confused classes are the denominations of 20 and 200 Mexican pesos and denominations of 50, 100, and 500 Mexican pesos, due to their similar colors. Table 2 shows the results of the training with the proposed regions: watermarks, portraits, symbols, denominations in number and words, and using the complete bill. (See Fig. 5b).

There is an overall efficiency of 47.24% and 80.54% with grayscale and color images, respectively.

Performance increased by providing more information on the content of the banknotes, but the majority of the incorrectly labeled banknotes were

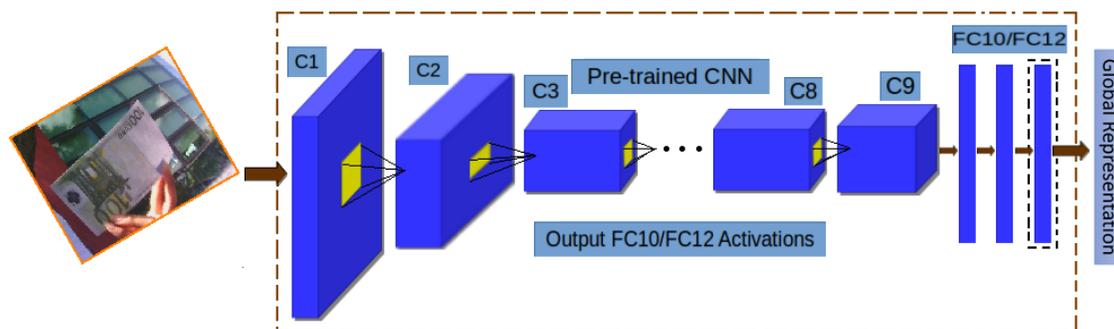


Fig. 7. Architecture of the pre-trained model of the Convolutional Neural Network. C1-C9 are convolutional layers and FC10/F12 are fully-connected layers. Inputs are images and outputs are global representations of each class object

Table 2. Proposing all objects on the banknotes

Mexican banknotes	Controlled		Semi-controlled		Not Controlled	
	gray	Color	gray	Color	gray	Color
20 pesos	72%	100%	8%	40%	5.83%	94.16%
50 pesos	92%	100%	56%	74%	0.83%	5.83%
100 pesos	96%	100%	100%	100%	90.83%	95.83%
200 pesos	50%	100%	42%	100%	9.16%	95.83%
500 pesos	32%	100%	54%	100%	0%	2.5%
Efficiency	68.4%	100%	52%	82.81%	21.33%	58.83%
Glb eff. gray	47.24%					
Glb eff. color	80.54%					

due to similarities in the silhouettes of the people, i.e., the silhouette has an upper body and a face, which causes the figures to look almost the same, so they are difficult for the network to differentiate.

Therefore, the portraits, watermarks, and symbols were eliminated as key features, resulting in a smaller set of bill regions. The result is in Table 3.

Table 3. Proposing the complete banknote and the denomination in number and words

Mexican banknotes	Controlled		Semi-controlled		Not Controlled	
	gray	Color	gray	Color	gray	Color
20 pesos	100%	100%	8%	72%	4.16%	95.83%
50 pesos	100%	100%	32%	88%	0%	10%
100 pesos	100%	100%	100%	100%	94.16%	99.163%
200 pesos	100%	100%	38%	100%	9.16%	97.5%
500 pesos	100%	100%	100%	100%	4.16%	5.83%
Efficiency	100%	100%	55.6%	92%	22.33%	61.67%
Glb eff. gray	59.31%					
Glb eff. color	84.56%					

There is a 59.31% overall efficiency with grayscale images and of 84.56% with color. This experiment has significantly reduced the confusion between classes; the percentage of error is due to an unclassified image, an image being very far from the camera, or the amount of light on the banknotes saturating the image. Once a set of discriminant characteristics in the banknotes was established, more samples were taken to train the network and thus increase efficiency. One thousand samples for each banknote denomination; for Mexican banknotes, the training set was composed of five thousand images and seven thousand images were used for euros.

Table 4. Increasing the number of samples to one thousand for each denomination

Mexican banknotes	Controlled	Semi-controlled	Not controlled
20 pesos	100%	98%	100%
50 pesos	100%	98%	91.67%
100 pesos	100%	100%	99.17%
200 pesos	100%	100%	100%
500 pesos	100%	100%	89.16%
Efficiency	100%	99.2%	96%
Glb eff.	98.4%		

This network had an overall efficiency of 98.4% for Mexican banknotes. Table 4 shows the results of this experiment. The set of tests for Mexican banknotes was performed with new samples that did not train the network. The image test set was 50, 50, and 120 controlled, semi-controlled, and uncontrolled, respectively for each banknote

denomination. From this experiment, it can be deduced that the key characteristics selected are suitable for a high recognition rate. The system has a better generalization with new samples.

In the case of the Mexican banknotes, the classifier here proposed takes both the color and the geometric features of the proposed regions into account. Generally speaking, a good training requires that the user provide data, even with redundancy, so that the algorithm can find the correlation between the variables. In the case of euros, the tests were performed on uncontrolled images, i.e., with variable distances. The training was carried out with one thousand samples of each denomination, with seven thousand in total, and 200 new samples per denomination were used for testing, totaling 1,400 samples. The results are shown in Table 5.

Table 5. Increasing the number of samples to one thousand for each denomination

Euros	
Denomination	Efficiency
5 euros	100%
10 euros	100%
20 euros	100%
50 euros	100%
100 euros	99.5%
200 euros	100%
500 euros	100%
Global efficiency	99.93%

Note that this method is still robust for Mexican peso and euro banknotes, as the algorithm focuses on the geometric and color features of certain proposed regions. Fig. 7 shows the predictions for euro banknotes.

7 Conclusion and Future Work

We have proposed a system for the recognition of banknote denominations with an impressive efficiency for Mexican peso and euro banknotes. This work considered a CNN model with deep learning for Mexican peso and euro. The model was trained with a supervised algorithm, requiring a high quantity of labeled data for training. One thousand pictures were considered for each banknote denomination in the training. The results

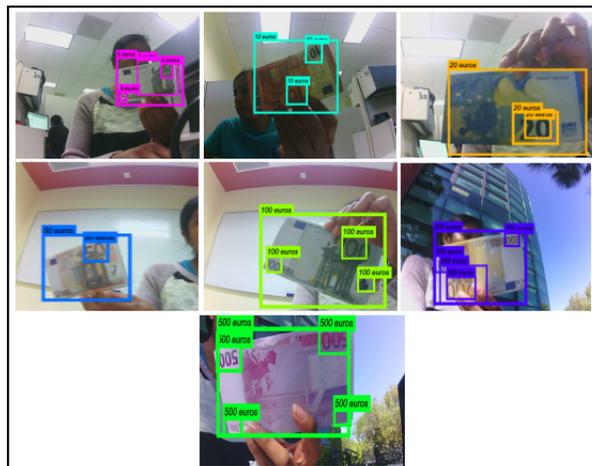


Fig. 8. Test images used for the recognition of euro banknotes

surpass other known works in efficiency. The color provides valuable information, so the network has a better performance and shows better results with the values of the banknotes written in words and numbers and with the complete banknote. As a future work, the database will be extended, adding American and Colombian banknotes and carrying out the classification between countries by applying the benefits of deep learning neural networks.

Acknowledgements

We would like to thank the National Council of Science and Technology (CONACYT, in its Spanish acronym) for the support received during the whole process of this work.

References

1. Ba, L. J. & Frey, B. (2013). Adaptive dropout for training deep neural networks. *Advances in Neural Information Processing Systems 26 (NIPS 2013)*.
2. Baldi, P. & Sadowski, P. (2013). Understanding dropout. *Advances in Neural Information Processing Systems 26 (NIPS 2013)*, pp. 1–9.

3. **Bebis, G., Boyle, R., Parvin, B., Koracin, D., Porikli, F., Skaff, S., Entezari, A., Min, J., Iwai, D., Sadagic, A., Scheidegger, C., & (eds.), T. I. (2016).** *Advances in Visual Computing: 12th International Symposium, ISVC 2016, Las Vegas, NV, USA, December 12-14, 2016, Proceedings, Part II*. Lecture Notes in Computer Science 10073. Springer International Publishing, 1 edition.
4. **Deshpande, A. (2015).** The 9 deep learning papers you need to know about. url-<https://adeshpande3.github.io/adeshpande3.github.io/> Link consulted 06-01-2018.
5. **Gai, S. (2016).** New banknote defect detection algorithm using quaternion wavelet transform. *Neurocomputing, Volumen 196*, pp. 133–139.
6. **Gai, S., Yang, G., Zhang, S., & Wan, M. (2013).** New banknote number recognition algorithm based on support vector machine. *IAPR Asian Conference on Pattern Recognition*, pp. 176–180.
7. **García-Lamont, F., Cervantes, J., & López, A. (2012).** Recognition of mexican banknotes via their color and texture features. *Trans. Expert Systems with Applications, Issue 10*, volume 39, pp. 9651–9660.
8. **Girshick, R. (2015).** Fast r-cnn. *IEEE International Conference on Computer Vision (ICCV)*, pp. 1440–1448.
9. **Goodfellow, I., Bengio, Y., & Courville, A. (2016).** *Deep Learning*, chapter Machine Learning Basics. MIT Press, pp. 98–155. Url-<http://www.deeplearningbook.org>.
10. **He, K., Zhang, X., Ren, S., & Sun, J. (2015).** Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *arXiv preprint arXiv:1502.01852v1*.
11. **He, K., Zhang, X., Ren, S., & Sun, J. (2016).** Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778.
12. **Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. R. (2012).** Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*.
13. **Jin, O., Qu, L., He, J., & Li, X. (2015).** Recognition of new and old banknotes based on smote and svm. *12th Intl Conf on Ubiquitous Intelligence and Computing and 12th Intl Conf on Autonomic and Trusted Computing and 15th Intl Conf on Scalable Computing and Communications (UIC-ATC-ScalCom)*, pp. 213–220.
14. **Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012).** Imagenet classification with deep convolutional networks. *International Conference on Neural Information Processing Systems*, volume 1, pp. 1097–1105.
15. **le Cun, Y. (1989).** Generalization and network design strategies. *Technical Report CRG-TR-89-4, University of Toronto*.
16. **Li, X., Zhang, G., Huang, H. H., Wang, Z., & Zheng, W. (2016).** Performance analysis of gpu-based convolutional neural networks. *45th International Conference on Parallel Processing (ICPP)*.
17. **n, S. G., Yang, G., & Wan, M. (2013).** Employing quaternion wavelet transform for banknote classification. *Neurocomputing, Volumen 118*, pp. 171–178.
18. **Pham, T. D., Kim, K. W., Kang, J. S., & Park, K. R. (2017).** Banknote recognition based on optimization of discriminative regions by genetic algorithm with one-dimensional visible-light line sensor. *Pattern Recognition*, volume 72, pp. 27–43.
19. **Redmon, J. (2013–2016).** Darknet: Open source neural networks in c. Link consulted 05-30-2018 <urlhttp://pjreddie.com/darknet/>.
20. **Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016).** You only look once: Unified, real-time object detection. *arXiv: 1506.02640v5*.
21. **Ren, S., He, K., Girshick, R., & Sun, J. (2016).** Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv:1506.01497v3*.
22. **Rojas-Domínguez, A., Lara-Alvarez, C., & Bayro-Corrochano, E. (2014).** Automated banknote identification method for the visually impaired. *CIARP*, pp. 572–579.
23. **Santos, L. A. (2016).** Artificial intelligence. <urlhttps://leonardoaraujosantos.gitbooks.io/artificial-intelligence/> Link consulted 05-31-2018.
24. **Simonyan, K. & Zisserman, A. (2015).** Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations (ICLR)*, pp. 1–14.
25. **Srivastava, N., Geoffrey Hinton, A. K., Sutskever, I., & Salakhutdinov, R. (2014).** Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research* 15, pp. 1929–1958.

26. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9.
27. Takeda, F., Nishikage, T., & Omatu, S. (1999). Banknote recognition by means of optimized masks, neural networks and genetic algorithms. *Engineering Applications of Artificial Intelligence*, pp. 175–184.
28. Tong, C., Lian, Y., Qi, J., & Xie, Z. (2017). A novel classification algorithm for new and used banknotes. *Mobile Networks and Applications, Issue 3*, volume 22, pp. 395–404.
29. Xie, L., Zhang, Z., He, S., Bo, X., & Song, X. (2017). Drug-target interaction prediction with a deep-learning-based model. *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 469–476.
30. Yeh, C.-Y., Su, W.-P., & Lee, S.-J. (2011). Employing multiple-kernel support vector machines for counterfeit banknote recognition. *Applied Soft Computing*, volume 11, pp. 1439–1447.
31. Youn, S., Choi, E., Baek, Y., & Lee, C. (2015). Efficient multi-currency classification of cis banknotes. *Neurocomputing, Volumen 156*, pp. 22–32.
32. Zeiler, M. D. & Fergus, R. (2014). Visualizing and understanding convolutional networks. *Computer Vision -ECCV, Part I*, pp. 818–833.
33. Zhang, X., Zhao, J., & LeCun., Y. (2015). Character-level convolutional networks for text classification. *In arXiv:1509.01626*.

Article received on 10/05/2018; accepted on 14/07/2018.
Corresponding author is Deysy Galeana-Pérez.