

Chapter of training machine learning models to automatically

[Business](#), [Industries](#)



CHAPTER ONE1. 1 Background1.

1. 1 IntroductionThe project proposed in this paper is essentially an Image Classifier for the different denominations of the Naira (Nigerian legal tender). Image classification is the process of training machine learning models to automatically group pictures into predefined classes utilizing low-level properties possessed by the images in question. 1 A decade ago, a project of this nature might only have been attempted at the postgraduate level but, thanks to recent advancements in image processing, computer vision, and artificial intelligence, accompanied with the democratization of machine learning tools² and techniques, today the completion of this project is feasible even at the undergraduate level. It would be important to note that this project is a part of a larger whole. The model proposed will play a critical role in developing a machine – similar to a cash deposit machine – capable of accepting a student’s card and crediting said card with the amount of cash deposited into the machine by the aforementioned student. The adoption of this machine will help usher Universities into the era of ‘ smart’ campuses. This will liberate staff and students from inconveniences associated with cash transactions and shift the existing paradigm towards cashless (digital) transactions using identification cards.

1. 1. 2 Problem StatementThe progress of today’s universities is stifled by cash-only transactions.

Staff and students alike are far too familiar with the inconveniences associated with cash transactions. Many a time a student will be unable to purchase a good or service, not due to insufficient funds, but instead due to

lack of smaller denominations of currency given to the student as balance. This project will play a crucial role in alleviating the aforementioned problem and in so doing expedite the emergence of ' smart' campuses.

1. 1. 3 Scope

The dataset for this classifier will be limited only to Naira notes. The implication of this is that if the model is fed an image of another currency or of a random object, the model will attempt to classify the said image into non-currency class, which is one of the predefined classes.

Also, the model being proposed assumes that all input images are legal tender and not counterfeit. This model is not intended to act as a counterfeit classifier – that on its own is a separate classification task. It would also be prudent to note that this project is purely software. Although, I will incorporate the model into the Intelligent Cashless Machine, which is also currently in development.

1. 2 Significance and Motivation

1. 2. 1

Significance

This project will contribute towards the proliferation of cashless, smart universities.

Also, the methodology chapter of this paper will shed light on current image classification and machine learning techniques; comparing neural networks and template matching.

1. 2. 2 Motivation

The motivation for pursuing this project stems from an innate technological curiosity. Pursuing this project allows me to explore the latest technology while, at the same time, solving a problem and nudging universities towards technological advancement.

1. 3 Aim and Objectives

1.

3. 1 Aim The aim of this project is to develop a machine learning based currency classification solution to aid the emergence of smart and cashless campuses. 1. 3. 2 Objectives i. To digitize all Nigerian currency notes as the dataset for this study.

ii. To design different Machine Learning (ML) models for unique identification of the digitized Nigerian currency notes. iii. To experimentally compare the ML models in (ii) based on appropriate performance metrics. iv.

To implement a Nigerian Currency Classification Application using the best model in (iii), which is capable of grouping a scanned Naira note into one of the eight different denominations of the Naira. v. To incorporate the classifier in (iv) into an Intelligent Cashless Machine (ICM), which can be deployed to achieve Smart and Cashless Campuses (SCC). 1.

4 Methodology A number of approaches were considered for the implementation of the model. A naive implementation was constructed using template matching techniques made available by the computer vision library; OpenCV. Currently, a more advanced implementation is being constructed using TensorFlow and a convoluted neural network. After much deliberation, I have decided to structure the project report as a review paper; noting the performance of both techniques (template matching and neural networks) and comparing their efficiencies in order to determine the more effective approach. 1. 5 Report Organization This report will follow the guidelines set by the Department of Electrical Electronics and Information Engineering, Covenant University. Below is an outline: 1. Chapter One –

Introduction This chapter provides the reader with a background of the project; highlighting the reasons for embarking on the project, the economic significance and its implementation.

2. Chapter Two – Literature Review This chapter documents a rigorous review of the relevant academic literature. It aims to provide the reader with an in-depth study of the advancements made by other academics in the field of image classification. This chapter concludes with a contribution to the field.

3. Chapter Three – Methodology This chapter details all information regarding the model design and the methodology employed. The chapter also poses a comparison between the various techniques employed in image classification.

It provides details on the implementation of the project with respect to the aims and objectives set for the project. 4. Chapter Four – Implementation and Testing This chapter contains the actual steps taken in the implementation of the project with diagrams, pictures, tests and results.

At this juncture the model should have reached a respectable level of accuracy and reliability. This chapter ensures against the release of a sub-standard product at the end of the project. 5.

Chapter Five – Conclusion This chapter delivers the final words concerning the project; highlighting observations, challenges, and recommendations for future researchers. CHAPTER TWO 2. 1 Introduction Image recognition, according to Ang et al. 3, is a field of technology that uses computers to

analyze, process and understand images so as to discern the target and object of different modes. In recent years, Image recognition has enjoyed increased attention from academia and industry. This is mostly due to the rise of computer vision smartphone applications, coupled with higher quality smartphone cameras⁴ and lower cost of higher performance microprocessors⁵. However, this progress is not just as a result of more powerful hardware; larger datasets, more robust models and new algorithms have contributed to the rise of image recognition applications⁶. Image classification, according to Bosch et al.

⁷, has to do with classifying an image by the object class that it contains. And, Das ⁸ defines classification as a machine learning task concerned with assigning labels to new data based on a given set of labeled data, referred to as the training set. Image classification is a product of image recognition, and both fields promise to play important roles in the future of artificial intelligence³ due to their many real-world applications.

2. 1. 1 Types of Classification There are two major approaches to solving classification problems with machine learning: supervised and unsupervised⁸.

According to Das ⁸, supervised classification uses spectral signatures gleaned from training data to classify an input image. In supervised classification, the training data set contains images labeled by a human. Machine learning practitioners take particular care in the building of their data sets. They make sure to introduce a degree of variance into their

training data set, so as to prevent the model from overfitting. Overfitting, according to Brownlee 9, is a phenomenon in machine learning where the model learns both the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. While, according to Das 8, unsupervised classification involves the discovery of spectral classes in an image without human intervention.

Ghahramani 10 posits that unsupervised learning should be thought of as finding patterns in data beyond what would be considered pure unstructured noise. The classification would then be carried out with the creation of clusters based on the patterns learned.

This project was carried out using supervised classification. The reasons behind this choice lie in the nature of the task at hand. Unsupervised classification is usually applied to large data sets, in which the machine learning practitioner is trying to glean novel insights. And seeing as how the task at hand is focused on teaching the model to recognize images of naira for later identification – not gleaning new information from said images – supervised classification is the appropriate choice. 2.

1. 2 Machine Learning Algorithms Used for Image Classification Quite a number of machine learning algorithms have been applied to image classification problems. Popular choices include: 1. Template Matching 2. Support Vector Machines (SVMs) 3. k-Nearest Neighbors (k-NN) 4. Hidden Markov Models (HMM) 5.

Deep Learning I will now provide a brief overview to each of the above-mentioned algorithms. 2. 1.

3 Template Matching Template matching techniques are an important ingredient of many computer vision systems ranging from image classification to quality control¹¹. But, the underlying principle behind template matching remains quite simple. The algorithm tries to find small parts of an image which match a template image, the template image is chosen after observing the unique patterns present in an image.

The algorithm is favored for its reliability and simplicity. But, ultimately, it is its simplicity that poses its greatest weakness. In template matching contexts are hardly constrained, as a result the slightest deviations in size, shape or orientation would prevent template matching algorithms from completing seemingly simple but nuanced recognition tasks. 2. 1. 4 Support Vector Machines (SVMs) Fradkin and Muchnik¹² propose that support vector machines be considered as a method for forming a special rule, called a linear classifier, such that classifiers with theoretical guarantees of good predictive performances are produced. Support vector machines are supervised learning models that rely on statistical learning algorithms to study data.

SVMs have been used in a wide range of applications; from facial expression recognition¹³ to gene analysis¹⁴. The foundations of support vector machines were developed by Vladimir N. Vapnik in 1995¹⁵. Support vector machines have been praised for their capability for generalization (a problem

traditional neural networks struggle with). The biggest limitation for SVMs, however, lies in the choice of kernel function parameters¹⁶.

1.5 k-Nearest Neighbors (k-NN) According to Altman¹⁷, the k-nearest neighbor algorithm is a simple non-parametric classification strategy in which the input consists of the k closest training examples in the feature space and new cases are classified based on a similarity measure (i. e. the object is classified based on the majority vote of its neighbors). The algorithm favored by data scientists because the cost of the learning process is zero, together with the fact that the algorithm makes no assumptions about the characteristics of the concepts required for the classification task. But, all in all, the algorithm falls short due to its high cost of computation, since the algorithm computes the distance of each query instance to all training samples¹⁹.

2.1.6 Hidden Markov Models (HMM) The mathematics behind the Hidden Markov model was developed in the 1960s by Baum et al.^{18, 22}. HMMs have gained popularity in recent years for their success in speech recognition²³. They are built on top of a basic Markov Chain²⁴; a stochastic model outlining a sequence of possible events in which the probability of each event depends only on the state reached in the previous event. According to Rabiner et al.

²⁵, a Hidden Markov model is a doubly stochastic process that is not observable but can only be observed through another set of stochastic processes that produce the sequence of observed symbols. Hidden Markov

models are lauded for their strong statistical foundation and efficient learning algorithms. But, some criticisms concerning HMMs often refer to the large number of unstructured parameters they require and the fact that they cannot express dependencies between hidden states. 2.

1. 7 Deep Learning Deep learning is a relatively new research area in machine learning 3. The technology is inspired by the natural neural networks present in the human brain. The deep learning architecture based on convolutional neural networks has achieved a sizable performance improvement on large-scale image classification tasks. Convolutional Neural Networks The convolutional neural network is a multi-layer network structure gained from the traditional artificial neural network. Artificial Neural Network is imitated from the human beings' observation of the principles of animals' neural system working, learning and memory approaches. Scientists first proposed neural network model known as Perceptron, by a number of input data a_1, a_2, \dots

\dots, a_n , corresponding to multiple weight values w_1, w_2, \dots

w_n , an output threshold t and an output value b . $b = \text{limits thing}^*$ from cat image classification paper Weights are initialized by manually giving the initial value, and adjusted in the process of iterating the training set, by specifying the increase in learning rate (learning rate) and constantly correcting the error (derived value and the actual does not match). Convolution and Pooling Convolution is common in computer vision, image processing and so on. Because of the ability to extract image

information, it is often used to extract image features in deep learning and image classification. Activation function In computational networks, the activation function of a node defines the output of that node given an input or set of inputs. A standard computer chip circuit can be seen as a digital network of activation functions that can be “ ON” (1) or “ OFF” (0), depending on input. This is similar to the behavior of the linear perceptron in neural networks. However, only nonlinear activation functions allow such networks to compute nontrivial problems using only a small number of nodes.

In artificial neural networks, this function is also called the transfer function. Popular choices: 1. Linear function 2. ReLU 3. tanh 4.

Softmax 5. Sigmoid Expandiate on Activation functions, show graphs and all.

And CITE Rectified Linear Unit: An activation function described by the equation below: It gives an output x if x is positive and 0 otherwise. The

Sigmoid function is an activation function described by the equation

below: 2. 2 Review of Related Literature I will now proceed to review a number of papers concerning image classification approached from different perspectives and implemented using various techniques. The aim of this section is to provide insight into the decisions made in the implementation of the model. And to justify said decisions by juxtaposing them with decisions made in recent academic studies.

In this vein, I have divided this section into three portions: 1. Review of Studies on Neural Networks for Image Classification 2. Review of Studies on

Other Methods for Image Classification
3. Review of Studies on Paper
Currency Recognition 2.

3 Review of Studies On Neural Networks For Image Classification
In this section, I will review studies in image classification implemented using neural networks.
2. 3. 1 Review of Tensorflow and Keras-based Convolutional Neural Network for Cat Image Recognition
In this study, Ang et al. 3 set out to build a supervised learning model capable of accurately identifying pictures containing cats. Two datasets were prepared in the development of the model; a training dataset and a test dataset. The model was trained on a dataset of 209 pictures.

Each image was given a fixed width and length, and each pixel was represented by an RGB value. The purpose of this was to enhance the performance of the convolutional neural network as it extracted and learned image features. The convolutional neural network was written in Python and implemented using Tensorflow and Keras. Ang et al. designed their convolutional neural network to consist of a two-layer convolution layer (each with a pooling layer), a two-layer feed-forward neural network and a 5×5 convolution kernel. The first convolution layer leads to a 32-layer convolution kernel.

And, the second convolution layer is used to derive a 64-layer convolution kernel. The convolution kernel is flattened through the reshaping process and imported into the 4096 feed-forward neurons. ReLU was used as the hidden layer activation function and Sigmoid as the final output activation

function. The model was trained a total of 40 iterations and when tested showcased a 90% accuracy rate, with a mean squared error of 0.29. Ang et al. concluded by outlining the current limitations of their model; like the fact that it can only work with the specified format of the static images.

And urged researchers to invest in further in-depth study of convolutional neural networks for image classification.

2.3.2 Review of ImageNet Classification with Deep Convolutional Neural Networks

In this paper, Krizhevsky et al. [26] share how they trained a large, deep convolutional neural network to classify the 1.2 million images in the ImageNet ILSVRC-2010 contest into 1000 different classes. And how they achieved by far one of the best results ever reported on the subsets of ImageNet used in the ILSVRC-2010 and ILSVRC-2012 contests.

As implied above, the dataset used in the training of the model is a subset of ImageNet with about 1000 images in each of the 1000 categories, summing up to roughly 1.2 million training images, 50,000 validation images, and 150,000 testing images. The dataset consisted of variable-resolution images. As a result, the images were down-sampled to a fixed resolution of 256 x 256, to accommodate their systems requirement of constant input dimensionality. They also subtracted the mean activity over the training set from each pixel, so that the network was trained on the raw RGB values of the pixels. The architecture of the network consisted of eight learned layers; five convolutional and three fully-connected.

When considering which activation function to use in the modeling of the neuron; Krizhevsky et al. opted for ReLU over tanh. Stating that deep convolutional neural networks with ReLUs train several times faster than those with tanh units. The model was trained on two GPUs, upon realizing that a single GPU with 3GB memory would be insufficient for the 1.2 million training examples. They spread the net across the two GPUs employing a cross-GPU parallelization scheme that puts half the kernels on each. They also decided to allow communication only in certain layers of the GPUs.

This was done to allow them to precisely tune the amount of communication until it is an acceptable fraction of the amount of computation. Krizhevsky et al. compared their resulting architecture to the “columnar” convolutional neural network employed by Cireşan et al. [27], the major difference being that the columns in their implementation are not independent. This scheme reduces our top-1 and top-5 error rates by 1.7% and 1.2%, respectively, as compared with a net with half as many kernels in each convolutional layer trained on one GPU.

The two-GPU net takes slightly less time to train than the one-GPU net [22].

4 Review of Studies On Other Methods For Image Classification

2.5 Review of Studies On Paper Currency Recognition

References and Bibliography

Chen and J. Z. Wang, “Image categorization by learning and reasoning with regions,” *Journal of Machine Learning Research*, vol. 5, pp. 913-939, 2004.

2) J. G. Tanenbaum, A. M.

Williams, A. Desjardins, and K. Tanenbaum, “ Democratizing technology: pleasure, utility and expressiveness in DIY and maker practice,” in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2013, pp. 2603-2612. 3L. Ang, Y.-x. Li, and X.

-h. Li, “ TensorFlow and Keras-based Convolutional Neural Network in CAT Image Recognition,” DEStech Transactions on Computer Science and Engineering, 2017. 4E. Koukoumidis, M. Martonosi, and L.-S.

Peh, “ Leveraging smartphone cameras for collaborative road advisories,” IEEE Transactions on mobile computing, vol. 11, pp. 707-723, 2012. 5S.

Furber, “ Microprocessors: the engines of the digital age,” in Proc.

R. Soc. A, 2017, p. 20160893. 6C.

Szegedy, W. Liu, Y. Jia, P.

Sermanet, S. Reed, D. Anguelov, et al., “ Going deeper with convolutions,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp.

1-9. 7A. Bosch, A. Zisserman, and X. Munoz, “ Image classification using random forests and ferns,” in Computer Vision, 2007. ICCV 2007.

IEEE 11th International Conference on, 2007, pp. 1-8. 8T. Das, “ Machine Learning algorithms for Image Classification of hand digits and face recognition dataset,” Machine Learning, vol. 4, 2017. 9J. Brownlee, “

Overfitting and Underfitting With Machine Learning Algorithms,” in Machine Learning Mastery, ed.

<https://machinelearningmastery.com>, 2016. 10Z. Ghahramani, “ Unsupervised learning,” in Advanced lectures on machine learning, ed: Springer, 2004, pp. 72-112.

11R. Brunelli, Template matching techniques in computer vision: theory and practice: John Wiley & Sons, 2009. 12D. Fradkin and I. Muchnik, “ Support vector machines for classification,” DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 70, pp. 13-20, 2006.

13P. Michel and R. El Kaliouby, “ Real time facial expression recognition in video using support vector machines,” in Proceedings of the 5th international conference on Multimodal interfaces, 2003, pp. 258-264. 14I. Guyon, J. Weston, S.

Barnhill, and V. Vapnik, “ Gene selection for cancer classification using support vector machines,” Machine learning, vol. 46, pp. 389-422, 2002.

15C. Cortes and V. Vapnik, “ Support-vector networks,” Machine learning, vol. 20, pp.

273-297, 1995. 16J. A. Suykens and J. Vandewalle, “ Least squares support vector machine classifiers,” Neural processing letters, vol. 9, pp. 293-300, 1999.

17N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, pp. 175-185, 1992. 18L. E.

Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *The annals of mathematical statistics*, vol.

41, pp. 164-171, 1970. 19L. E.

Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, p. 1883, 2009. 20L. E. Baum and J.

A. Eagon, "An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology," *Bulletin of the American Mathematical Society*, vol. 73, pp. 360-363, 1967.

21L. E. Baum and G. Sell, "Growth transformations for functions on manifolds," *Pacific Journal of Mathematics*, vol.

27, pp. 211-227, 1968. 22L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state Markov chains," *The annals of mathematical statistics*, vol. 37, pp. 1554-1563, 1966.

23D. B. Paul, "Speech recognition using hidden markov models," *The Lincoln Laboratory Journal*, vol. 3, pp. 41-62, 1990. 24A.

A. Markov, “ An example of statistical investigation of the text Eugene Onegin concerning the connection of samples in chains,” *Science in Context*, vol. 19, pp. 591-600, 2006.

25L. Rabiner and B. Juang, “ An introduction to hidden Markov models,” *IEEE ASSP Magazine*, vol.

3, pp. 4-16, 1986. 26A.

Krizhevsky, I. Sutskever, and G. E. Hinton, “ Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 2012, pp. 1097-1105. 27D. C.

Cire? an, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, “ High-performance neural networks for visual object classification,” *arXiv preprint arXiv: 1102.*

0183, 2011.