Bhutanese currency recognition using Convolutional Neural Network

Parshu Ram Dhungyel[1*], Manoj Chhetri[2]

Information Technology Department, Collage of Science and Technology, Royal University of Bhutan.
parshuram.cst@rub.edu.bt[1*], manoj_chhetri.cst@rub.edu.bt[2]

Abstract

Currency recognition, or digitization, refers to the process of converting physical currency, such as banknotes and coins, into digital formats. This transformation enhances convenience, security, accessibility, and cost efficiency in financial transactions, which are essential in the modern economy. This paper introduces a Convolutional Neural Network (CNN) model designed for the recognition of Bhutanese paper currency. The model was trained on a dataset comprising various currency types and denominations, achieving a training accuracy of 91 percent and a testing accuracy of 80.5 percent. The architecture consists of three convolutional layers, followed by a dense layer for classification. The findings suggest that CNNs are effective for currency recognition, with the potential for improved accuracy through the expansion of the training dataset.

Key Words: *Bhutanese currency, Currency recognition, Convolutional Neural Network (CNN), Image processing.*

1. INTRODUCTION

1.1. Background on the Bhutanese Currency

Bhutan introduced its first paper currency in 1974, with denominations of 1, 2, 5, 10, and 100 Ngultrum. The Royal Monetary Authority (RMA), established as the central bank in 1982, was granted the authority to issue currency in 1983. In 1986, the RMA introduced its own banknotes, which were later replaced by the current series of denominations ranging from one to one thousand Ngultrum, which have been in circulation since 2006.

Despite advancements in technology and the development of digital financial services such as m-BoB, a significant proportion of financial transactions in Bhutan continue to rely on paper currency. This reliance underscores the need for automation in cash handling, particularly in banks, retail establishments, vending machines, ATMs, and other financial institutions. Currently, self-service banking terminals in Bhutan can recognize card authorization and dispense cash based on predefined denomination values stored in separate compartments. However, for any financial automation system involving paper currency, an effective currency recognition system is essential to ensure efficiency and accuracy in financial operations.

1.2. Currency Identification approach

The present study aims to address the challenge of recognizing Bhutanese currency through the application of Convolutional Neural Networks (CNNs). Researchers in other countries have relatively easy access to extensive databases containing images of local currencies. However, there are no publicly available datasets specifically related to Bhutanese currency. The development of models or classifiers for recognizing Bhutanese currency and its respective denominations requires a substantial dataset comprising images of banknotes captured under varying conditions, including differences in age and lighting.

This study not only proposes a CNN-based approach for the identification of Bhutanese currency notes but also contributes to the creation of a comprehensive dataset. This dataset may serve as a valuable resource for future research on banknote recognition, facilitating further advancements in the field.

Fig. 1: *Images of Bhutanese currency*

2. RELATED WORK

Banu et al. (2022) present a method for counterfeit currency detection using MATLAB, with a focus on image acquisition and segmentation techniques. By employing Canny's algorithm for edge detection and feature extraction, their approach effectively differentiates counterfeit currency, particularly the newer 500- and 2000-rupee denominations. Similarly, Mittal and Mittal (2018) conducted a study on the recognition of Indian banknotes, utilizing deep learning-based classifiers in conjunction with transfer learning to identify currency denominations. Their dataset was collected under varying conditions, considering factors such as light intensity, posture, and the quality of the banknotes.

Several researchers have explored multi-currency recognition systems for regions where multiple currencies are in circulation. These systems first identify the country of origin before classifying the denomination using texture and color features, employing the k-nearest neighbor (KNN) algorithm. The extracted features are then compared with those of authentic banknotes using a convolutional neural network (CNN) (Chowdhury, 2020).

Although the primary objective of this study is to develop a system for recognizing Bhutanese currency, significant research has been conducted on counterfeit detection. Murthy (2016) identified an increase in counterfeit currency circulation and concluded that distinguishing counterfeit notes with the naked eye is challenging. However, machine learning models trained using CNNs demonstrate high accuracy in counterfeit detection.

Banknote identification remains a widely researched area due to its broad applications, including vending machines, currency recognition for visually impaired individuals, and Automated Teller Machines (ATMs). Numerous studies have employed traditional digital image processing techniques, manually extracting distinct features for currency recognition. For instance, Kamal et al. (2015) proposed a modular approach for identifying Indian banknotes by extracting unique features such as the central numeral, color band, RBI seal, and tactile marks for visually impaired individuals.

3. METHODOLOGY

The flowchart in Fig. 2 presents the process of creating a CNN. The Fig. 3 depicts a cyclical process of creating, training, testing, and validating a CNN model to achieve optimal performance in image recognition tasks, the steps are discussed in details below:

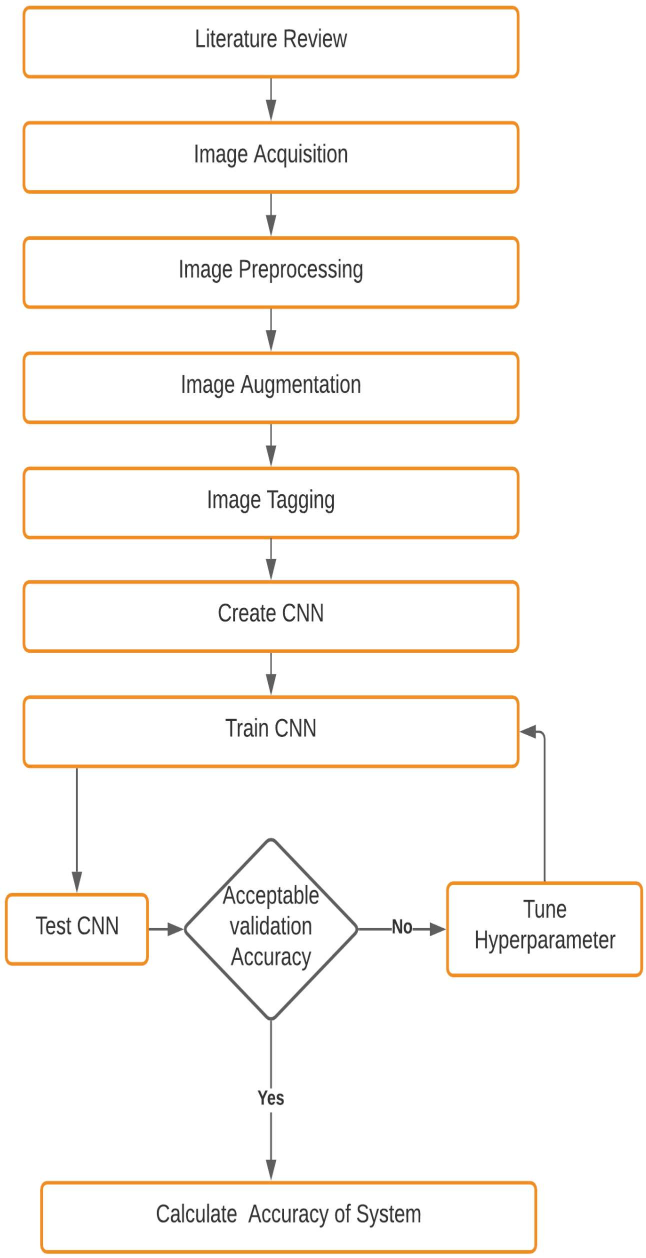

***Fig. 2:** Methodology for Currency recognition using CNN*

3.1. Approach/Paradigm Theory

The primary objective of this research is to develop a computational approach for recognizing Bhutanese currency. This study falls within the domain of computer vision, which involves enabling computer systems to interpret and analyze visual data.

In this study, the system is designed to first determine whether an input image represents Bhutanese currency. If identified as Bhutanese currency, the system subsequently classifies the denomination of the note. To achieve this, a deep learning-based neural network, specifically a CNN, was utilized. While alternative algorithms exist for similar applications, CNNs were selected due to their ability to automatically extract distinguishing features from images, eliminating the need for manual feature engineering. The CNN was designed to recognize a sufficient number of features to accurately differentiate between various denominations.

For instance, relying solely on color as a distinguishing feature may lead to misclassification, as multiple denominations can share similar color patterns. Additionally, banknotes may change in appearance over time due to aging or wear, altering their contrast and texture. For example, the Nu. 1000 note is predominantly yellow, and a model based solely on color differentiation might incorrectly classify all yellow-colored images as Nu. 1000. Thus, it is essential for the network to learn multiple features to enhance accuracy in currency identification.

Although a detailed methodology is presented in Fig. 2, the proposed approach can be broadly categorized into two key tasks:

- Dataset creation
- Training the CNN model on the dataset

The initial four stages in Fig. 2, following the literature review, focus on dataset creation, while the subsequent stages involve training and evaluating the CNN model. The literature review, highlighted as the first task, is an integral component of the research and is conducted throughout various stages of the study.

Dataset creation

The initial task undertaken in this research was the collection and preparation of a comprehensive dataset. Deep learning models, such as CNN, require a large volume of training data and are prone to overfitting, wherein the model memorizes the training images rather than generalizing to unseen data. While overfitting can be mitigated through programmatic techniques during model training, ensuring a sufficiently diverse and representative dataset is essential for enhancing model performance.

a. Image Acquisition

The primary dataset for this study was obtained through direct image capture using smartphones. Additionally, supplementary images were sourced from online repositories to enhance dataset diversity. The data collection process is outlined in detail in Section 3.1.

As previously discussed, the research team, along with student participants, was actively involved in data collection. The participation of students aimed to provide them with hands-on experience in data acquisition within the context of machine learning. Specifically, students from the third-year Electronics and Communication Engineering (ECE) and Information Technology (IT) programs enrolled in the CTE309 Multimedia Technologies course (Spring Semester 2022) were tasked with capturing images of Bhutanese currency notes and uploading them to a sshared Google Drive. Given that the course was conducted online at that time, students located in different dzongkhags had access to various currency denominations, contributing to dataset variability.

Furthermore, an additional dataset was collected during the Winter Semester of 2021, further increasing the diversity of the images used in this study.

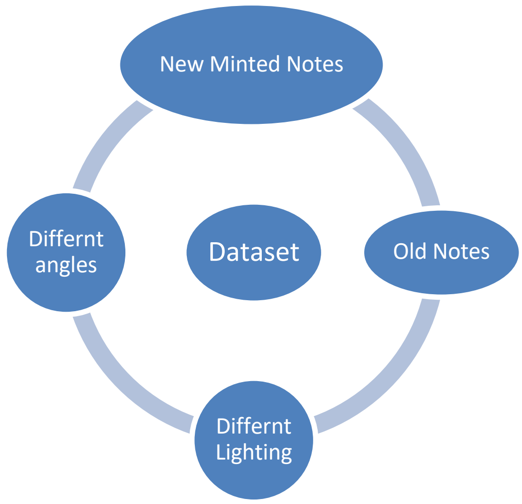

Fig. 3: *Image acquisition process in a nutshell*

3.2. Image Pre-processing

With advancements in deep learning, extensive pre-processing of images is no longer as critical as it was in earlier approaches. However, certain preprocessing steps remain essential to ensure the quality and consistency of the dataset. In this study, the preprocessing stage primarily involved assessing image quality, discarding unusable images, and cropping images to a standardized size. This process was meticulous and time-consuming, as many images were improperly captured, containing extraneous objects or lacking sufficient clarity. Ensuring that only high-quality images were retained was crucial for improving the model's ability to accurately recognize and classify Bhutanese currency.

Fig. 4: *Image before pre-processing*

Fig. 5: *Example of image after removing whitespaces*

3.3. Image Augmentation

For each image, multiple copies will be created by performing various transformation tasks like zooming (at different values), rotating, and shearing. The augmentation process was implemented using the ImageDataGenerator function in keras library.

While an increased number of augmented images generally improves model performance, the extent of augmentation was constrained by the computational resources available. Therefore, the number of augmentations was adjusted dynamically based on hardware limitations and the performance of the CNN model to optimize efficiency without compromising accuracy.

a. Image Tagging

As CNNs operate through supervised learning, it was necessary to label each image with its corresponding class or denomination value. This labeling process is labor-intensive, and although automatic tagging tools are available, they do not match the accuracy of manual tagging. Therefore, manual tagging was employed for this study. To facilitate efficient management, images were organized into folders, with each folder representing a distinct class (denomination value). During the testing phase, the class of each image could be determined based on the folder in which it was stored.

After labeling, the dataset was divided into training and testing subsets. Eighty percent of the images were allocated for training the model, while the remaining twenty percent were used for testing. While there is no universally standardized ratio for dataset splitting, adjustments to the training and testing proportions may be made based on the model's performance. For example, in cases of overfitting or underfitting, the dataset split may be modified to a 70:30 ratio or additional data may be collected. It is common practice to allocate a larger proportion of the data for training, as neural networks, which operate on a "learning by example" principle, perform better when provided with more data to learn from. The final test set was exclusively used for model validation, serving to assess the model's ability to accurately classify Bhutanese currency denominations.

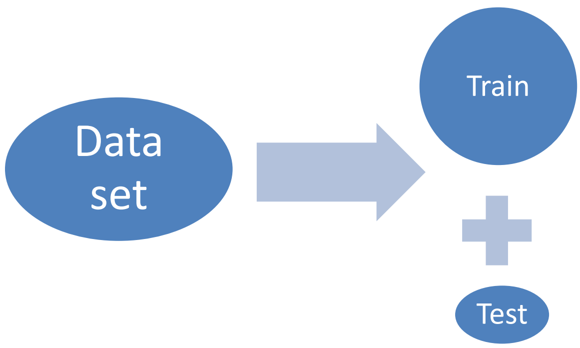

Fig. 6: *Dataset split into training and testing set*

3.4. Training and Testing the Model

a. Create CNN

The model is designed to accurately identify all relevant features of the currency. The CNN architecture primarily relies on convolution operations, utilizing multiple kernels to generate distinct feature maps. A Max Pooling layer is incorporated to retain only the most prominent pixels, ensuring that the most salient features are emphasized. Given that the convolutional process highlights particular features, the Max Pooling layer effectively reduces the spatial dimensions of the feature maps, allowing the network to focus on the most significant information.

Subsequently, a Flatten layer is employed to convert the two-dimensional feature maps into a one-dimensional vector. This transformation enables the vector to be passed to the Dense layer, which performs the final classification task.

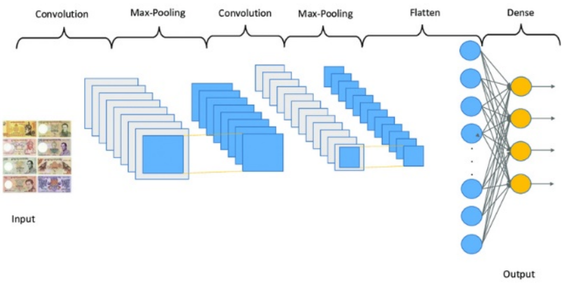

Fig. 7: *Working of CNN*

Artificial Neural Networks (ANNs), including Convolutional Neural Networks (CNNs), are often referred to as "black box" models due to the inherent difficulty in explaining the precise reasoning behind their predictions. While the internal operations of these networks remain complex, experienced machine learning engineers can often infer their underlying mechanisms. In the context of this study, for instance, it is known that one convolution layer may be responsible for detecting edge features, while another may focus on identifying color patterns, denomination numbers, and other distinctive characteristics.

To develop the CNN model, we initially constructed a simplified architecture by studying the known features of Bhutanese currency, as outlined in the literature. However, given the complexity of CNNs, it is not feasible to account for all potential features through literature alone. CNNs, as complex systems, possess the ability to detect novel patterns or features within image files (binary files or pixel matrices) that may not be perceptible to the human eye. Consequently, in addition to the layers designed to identify predefined features, we incorporated extra layers within the CNN to allow the model to autonomously learn and identify additional patterns. This adaptability is one of the key strengths of CNNs, enabling them to evolve as they process more diverse data.

b. Train CNN

The section of the image dataset designated for training the neural network will be further split into training and validation sets, typically following 80/20 or 70/30 ratio. A larger portion of the dataset was allocated for training the network. During the training process, the images are presented to the CNN multiple times, determined by the number of epochs, which generally ranges from 10 to several thousand. An epoch refers to one complete pass of the entire dataset through the network, and setting a higher number of epochs ensures that the CNN learns from the data more thoroughly. For instance, a 10-epoch configuration means the network will process each image ten times. Given that the images were carefully collected by the researchers, ensuring the accuracy and relevance of the data is critical, as erroneous data could result in inaccurate predictions. Therefore, meticulous attention was devoted to data preparation to guarantee that only correctly labeled and high-quality images are fed into the network, as this directly influences the model's performance.

c. Test CNN

After training the network, it is evaluated using the validation set to assess its performance. At this stage, issues related to overfitting and underfitting are addressed. If the validation results are suboptimal, the underlying causes are analyzed, and network parameters such as the number of convolutional layers, max-pooling layers, and the choice of optimizer are adjusted accordingly.

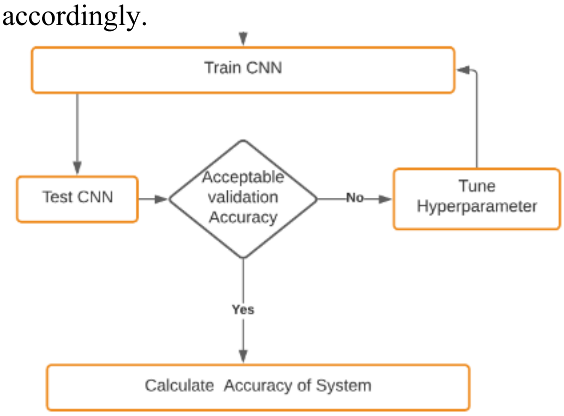

Fig. 8: *Training and Testing occurs multiple times until an acceptable accuracy is reached*

This iterative process of refining network parameters is known as hyperparameter tuning. Each modification to the network parameters necessitates retraining the CNN, requiring a return to the training phase. This cycle of adjusting hyperparameters and retraining continues until the model achieves an optimal performance on the validation set, ensuring a well-generalized and robust network capable of accurate currency recognition.

4. RESULT & DISCUSSION

4.1. Evaluation Metrics

The predictive performance of the CNN model in identifying Bhutanese currency was evaluated using accuracy as the primary metric. Initially, the network was trained on the training dataset and subsequently tested using the validation dataset. Once a satisfactory accuracy score was achieved on the validation set, the model was then tested on the test dataset. The accuracy obtained on the test dataset represents the final performance of the model. Further analysis was conducted to evaluate the model's ability to correctly identify the different denominations of Bhutanese currency. This was achieved by generating a confusion matrix, which provided insights into the specific denominations that the CNN model successfully recognized, as well as those it struggled to identify. The confusion matrix enabled a detailed examination of the model's strengths and weaknesses in denomination classification.

$$Accuracy = \frac{Correct\ Predictions}{All\ Predictions} = \frac{TP + TN}{TP + TN + FP + FN}$$

Fig. 9: *Accuracy formula*

TP: True Positive
TN: True Negative
FP: False Positive
FN: False Negative

4.2. Calculate accuracy of system

The performance of the model was evaluated using the 20% of the data reserved as the testing set, which had not been previously encountered by the network. The accuracy score obtained on the test dataset was considered as the final performance measure for the model. The goal was to achieve an accuracy of at least 80% with the system. The models were trained for a maximum of 1000 epochs, with early stopping applied and a patience value of 10 to prevent the model from running for unproductive epochs.
The final results indicated an accuracy of 91% on the training dataset and 80.5% on the test dataset. The model architecture comprised three convolutional layers with 128, 64, and 32 neurons, respectively. Each convolutional layer employed 32 filters with a kernel size of 3x3. A max pooling layer was introduced after each convolutional layer to reduce the spatial dimensions of the feature maps. Adam optimizer was used as the default optimization algorithm for training the network. These results demonstrate that the model achieved satisfactory performance in recognizing Bhutanese currency denominations, with the training accuracy surpassing the target accuracy threshold and the test accuracy meeting the expected standard.

5. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to the university research committee for providing the opportunity to conduct this research. Our appreciation also extends to the Dean of Research and Industrial Linkages and the Research Officer for their invaluable assistance and guidance throughout the research process. Additionally, we would like to thank the reviewers at both the college and university levels for their constructive feedback, which was instrumental in refining the focus of this study. We are also grateful to the students of the Third Year BE IT and ECE programme, whose active involvement in data collection significantly contributed to the success of this research project.

6. CONCLUSION

In conclusion, this study highlights the promising potential of using convolutional neural networks (CNNs) for the recognition and digitization of Bhutanese currency. The CNN model, trained on a substantial dataset of Bhutanese banknotes, demonstrated a high level of accuracy in classifying currency denominations.
The application of CNNs facilitated effective feature extraction and classification, making it a suitable and efficient approach for currency recognition tasks. The model achieved a training accuracy of 91% and a testing accuracy of 80.5%, indicating that CNNs can be a viable tool for developing systems capable of reliably identifying Bhutanese paper currency. These results suggest that the performance of the model could be further enhanced by increasing the size and diversity of the training dataset.
Ultimately, the digitization of currency holds significant potential to improve the convenience, security, accessibility, and cost-efficiency of financial transactions, positioning it as a critical element in the modern economy. This research

also underscores the applicability of CNNs in real-world problems, particularly in currency recognition. Future research could focus on refining the model to improve its robustness, particularly in handling counterfeit notes and variations in the physical appearance of banknotes over time.

REFERENCES

Jayant Kumar Nayak, Chaitan Majhi, Apurva Kumar Srivastav, & Ajaya Kumar Dash. (2015). *Neural network approach for Indian currency recognition*. https://doi.org/10.1109/indicon.2015.7443289

Mittal, S., & Mittal, S. (2018). Indian Banknote Recognition using Convolutional Neural Network. *2018 3rd International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU)*. https://doi.org/10.1109/iot-siu.2018.8519888

Kamal, S., Simarpreet Singh Chawla, Goel, N., & Raman, B. (2015). *Feature extraction and identification of Indian currency notes*. https://doi.org/10.1109/ncvpripg.2015.7490005

Murthy, S., Jayanta Kurumathur, & B. Roja Reddy. (2016). *Design and implementation of paper currency recognition with counterfeit detection*. https://doi.org/10.1109/get.2016.7916838

Chowdhury, U. R., Jana, S., & Parekh, R. (2020, March 1). *Automated System for Indian Banknote Recognition using Image Processing and Deep Learning*. IEEE Xplore. https://doi.org/10.1109/ICCSEA49143.2020.9132850

Sabitha Banu, A., & Padmavathi, D. G. (2022). Hybrid detection and mitigation of DNS protocol MITM attack based on Firefly algorithm with Elliptical Curve Cryptography. *EAI Endorsed Transactions on Pervasive Health and Technology, 8*(4), e3. https://doi.org/10.4108/eetpht.v8i4.308