Project Proposal
Universal Currency Recognition for the Visually Impaired

Rajkumar Rajavel, Khyati Jariwala, Nicolas Rey, Yu Xia, Raghav Vaidya

Problem Description

There is a large set of people that have visual impairments. Some of these visual impairments include having low visibility, partial blindness, complete blindness, and many others. Even though there are many people that fall within this category, the world is not very visual-impairment friendly. One domain that people with visual impairments face challenges with is recognizing different banknotes. This is a very important problem considering that money is involved in many daily transactions and this would likely hinder one's independence.

There are a few existing solutions that attempt to address the issue of currency recognition, however, they have limitations due to small datasets and limited currency coverage. Given this critical issue of currency recognition, we aim to develop a predictive modeling solution using a large dataset for universal currency recognition to assist individuals with visual impairments. Doing this would allow us to empower people with visual impairments to take control of their lives by enhancing accessibility and autonomy.

Data Description

The BankNote-Net dataset serves as a valuable resource for creating a predictive modeling solution for currency recognition. Since the goal is to predict the currency denomination and currency type, our dataset works wonderfully because it is a comprehensive dataset of banknote images, including multiple currencies and denominations. It consists of 24,816 embeddings of banknote images, which span 17 currencies and 112 denominations. The data was captured in various assistive scenarios, making it a diverse and realistic pool of data applicable to various real-world scenarios. Additionally, given that the dataset captures many currencies and denominations, we incorporate geographic diversity, which helps us access a larger number of potential visually impaired users.

Data Pre-Processing

To ensure that the dataset is prepared prior to beginning the modeling and exploratory data analysis, we will take some precautionary measures. This includes validating and standardizing the collected data to ensure consistency and quality. We will ensure this by ensuring that images will be resized to a common format (e.g., 224x224 pixels) for model compatibility. Additionally, any inconsistencies in labeling will be addressed during preprocessing. Some of these inconsistencies could be data cleaning (getting rid of outliers or bad images), addressing missing values (potentially through imputation), and even normalizing (making sure pixel values are on a common scale) as needed.

Exploratory Data Analysis (EDA)
Prior to jumping straight into running models for the predictive modeling necessary for currency recognition, we will conduct exploratory data analysis (EDA). This will allow us to conduct exploratory data analysis to understand the characteristics of the dataset. It will also help us identify potential challenges and opportunities within the dataset. Potential ideas for EDA with this particular dataset include analyzing image quality, variations in backgrounds, orientations, and other relevant factors. This will help inform future steps of feature engineering and model development.

Feature Engineering
Feature engineering will be crucial for extracting relevant information from banknote images. We will use supervised contrastive learning to effectively create embeddings that represent banknote images. These embeddings will serve as features for our predictive modeling. We may consider using convolutional neural networks (CNNs), especially pre-trained CNNs, to extract these high-level features from the banknote images. Additionally, there are other methods like local binary patterns (LBP) and histograms of oriented gradients (HOG) that can be used to capture textures and patterns within the banknote images.

Data Splitting
The dataset will be divided into training and testing sets, ensuring data separation for model evaluation. We will consider full splits (80%-20%) for abundant data settings and few-shot splits (10%-90%) for limited data scenarios.

Model Selection
We will explore various machine learning models and architectures, considering their suitability for universal currency recognition. Models will be selected based on their potential to handle multi-class classification across different currencies and denominations. Some examples of models we are considering using for the predictive modeling problem are convolutional neural networks, given how well they perform for image-based tasks, and transfer learning, which allows us to utilize pre-trained models to save time and resources for the purposes of currency recognition.

Model Training and Evaluation
The model that we select will be trained using the dataset and features generated through supervised contrastive learning. Hyperparameter tuning will be performed to optimize model performance and avoid issues related to overfitting. We will also explore

three different initialization scenarios: BankNote-Net, ImageNet, and Random to compare model performance and select the best one. Loss curves can help with detecting if the model is overfitting or underfitting and assess the overall convergence of the model. Once the model is trained, we will evaluate the performance of the model using metrics like accuracy, precision, recall, F1 score, and confusion matrices. The confusion matrix will be a very useful visualization for this purpose and one can be made for each currency denomination and type. Receiver operating characteristic (ROC) curves and area under the curve (AUC) can also be used if our approach to the predictive modeling problem is binary in nature (i.e. 1 = currency correctly recognized, 0 = currency not correctly recognized).

Results

The predictive model developed in this project for the purposes of currency recognition will aim to achieve high accuracy, recall, and precision in recognizing banknotes. The results of the model will be presented in terms of model performance for both full and few-shot scenarios accompanied by key visualizations that will be elaborated on for analysis.

Analysis

We will analyze the strengths and weaknesses of the selected model and its performance on different currencies. One way of doing this is to create class activation maps (CAM) or gradient-weighted class activation maps (Grad-CAM). This technique helps us visualize what regions of the banknote image inputs contribute the most to the model's prediction power and informs us on the key parts of a banknote image that are used for classification. Also, we could use t-SNE or PCA embeddings, which reduce the dimensionality of feature vectors, to visualize features and assess how well the model separates different currency denominations or types.

We will also assess the impact of embeddings derived from supervised contrastive learning. A discussion of the findings will be included, potentially focusing on the generalization of the model to unseen currencies. Using our results, we will elaborate on future steps that could be taken to improve and better the lives of visually impaired people when it comes to currency recognition.

Preliminary References

1. https://paperswithcode.com/dataset/banknote-net (dataset)
2. https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment (global stats on blindness)
3. https://www.afb.org/blindness-and-low-vision/using-technology/accessible-identification-systems-people-who-are-blind-0 (existing solutions and use case)