

## **Open CV based Information Extraction from Cheques**

**Aniket Dhanawade, Abhishek Drode, Gifty Johnson, Aadesh Rao and Dr. Savitha Upadhya**

The traditional verification process of bank cheques includes date, signature, legal information, and payment written on the cheques. In this paper, extracting the legal information from captured cheque image is obtained by pre-processing the image, extracting required information, then recognizing and verifying the handwritten fields. Image processing techniques like thinning, median filtering, dilation, and verification techniques have been used in this paper.

### **POINTERS –**

- 1) All image processing techniques are implemented and used using OCR, tesseract and OpenCV.
- 2) The recognition system takes a scanned image as an input image. The image should have a specific format as JPEG. This image is taken through a scanner or a camera.
- 3) All cheques have a standard format. The date, payee, signature, amount is all in the same place. So manual cropping is performed i.e. the coordinates of the image to be cropped is given manually to extract them.
- 4) Binary thresholding before passing it to the OCR. Linear normalization is performed using the given formula: 
$$In = (I - Min) * (newMax - newMin) / (Max - Min) + newMin$$
 Where 'I' = grayscale image with intensity values ranging from (Min, Max) and 'In' = New image with intensity values ranging from (newMin, newMax).
- 5) The text from the cropped image is now read using tesseract library in Python that perform OCR on the images and extracts the text from them. All these are saved in a csv file.
- 6) After extracting the cheque features such as the bank name, cheque amount, account number, bank address and the date, the next step is to authenticate the signature which can be done using Neural Network or Machine Learning algorithms.

### **CONCLUSION –**

The objective of the above work presented in this paper is to make an automated system which will extract information from cheques, validate the information and keep record of all the information present that can be used by financial institutes to reduce human efforts in verification of signature which is done manually as well as it will keep a record of all the transactions. In this paper bank name, address, account number, date, amount was extracted from the cheque using different libraries in Python such as OpenCV and tesseract. By using Pandas, the data obtained was stored in the form of excel sheet.