An Image-based Intelligent System for Data Extraction

Shawn Louis
Department of Computer Engineering,
Don Bosco Institute of Technology,
Mumbai, India
shawnlouis2000@gmail.com

Piyush Sonar
Department of Computer Engineering,
Don Bosco Institute of Technology,
Mumbai, India
piyushdsonar@gmail.com

Priya Kaul
Department of Computer Engineering,
Don Bosco Institute of Technology,
Mumbai, India
priya.dbit@dbclmumbai.org

***Abstract*—Automation is the process of providing goods and services with fewer to no human interventions. The major advantage of automation is reduction in human error. The system proposes to extract data from images that are tilted at different angles and noisy. The system reduces human error by storing the data directly in the database. The proposed system will take image input from the user through a user interface. This interface is a web application. The input image is pre-processed and forwarded to a machine learning model. The machine learning model is trained and tested using a character data set and convolutional neural network. The model will detect the characters and will give the output as recognized text. This output will be automatically stored in the database and shared with the user through the same interface.**

Keywords—Image processing, Machine Learning, Neural Network

I. Introduction

There has been enormous change in technology over the past few years. With new advancements in technology, there has been an increase in the size of data generated. Besides this, lots of data is generated through social media sites, video streaming, and applications. People use lots of appliances and products which generate data in large amounts. This data is very difficult to handle manually. Fortunately, we have different options such as hardware storage, cloud storage for data storage purposes. The difficult task is to automate a process that can operate data and data operations.

Automation is the process of providing goods and services with fewer to no human interventions. The major advantage of automation is the reduction in human labor. This eventually reduces human error. For instance, an automated machine that sorts products into different categories. For humans, there is a possibility of vision error. Human error can occur at times of classifying the product into a specific category. In the case of an automated machine, this error can be reduced. Also, an automated machine will take less time than a human to separate products. So automation saves time and reduces human labor.

When a company expert goes to a customer's house to take gas and electricity meter readings, it becomes very difficult for him to take a picture of the meter and simultaneously store the reading into the application. An intelligent data extraction system will automatically extract data from the meter picture and store the data directly in the database. This will save human labor and increase the efficiency of the meter reading system.

The intelligent system can be used for data extraction from documents, car number plates, and vehicle chassis. This system will be useful where the input image may contain background noise and be tilted at some angle. The existing system provides a charged service for data extraction. This project aims to provide an open-source solution for data extraction from different input images.

II. Related Work

Analysis of various research papers has been carried out. Each paper talks about different strategies for data extraction. They also explain various image processing algorithms required before data extraction. The image preprocessing is carried out to enhance the image quality before extracting the data. Different algorithms are used along with available image pre-processing methods to get rid of background noise from the image. Edge detection methods are used to determine the characters present in the image. For better accuracy, neural networks can be used for training the data extraction model.

Geetha et al. [1] developed a piece of software that uses OpenCV and Nanonet OCR to extract text from formatted bills. CNN is being used to carry out this endeavour. Using various algorithms and software, we were able to recognise text in structured money. The experiment shows that the new procedure is more accurate and efficient than the old one. It's also possible to find methods for extracting text from unformatted bills and updating it into the database automatically. This project's purpose is to investigate the task of converting handwritten text and invoices into an excel format. Invoices are frequently processed manually. As a result, automation is used in our system, potentially reducing manual labour. Nanonet automatically processes bills based on trained data by evaluating the content using machine learning algorithms.

Duan et al. [2] present a method that works well with a variety of VLP images, including scratched and scaled plate images When working with low-quality plates, however, it still has a few flaws. The findings would be better if we utilised a more contemporary camera, and we would be able to remove numerous inaccuracies caused by faulty plates. For example, we can capture high-quality photos of plates that are obscured by muck or dust using an infrared camera. In conclusion, combining the Hough transform with the contour algorithm improves the accuracy and speed of VLP detection. As a result, the method might be used in real-time systems. This approach is used in our automatic VLPs recognition system in practise.

Kanagarathinam et al. [3] Because of the discontinuity in the digit representation, existing open-source OCR software could not read the text of seven-segment numerals. The use of OCR could aid in the automation of the energy management process. Their dataset can also be used to automate the usage of gas and water. The research also looks at how a smart metering system for measuring and managing electricity demand might help reduce infrastructure costs. In comparison to existing methods, the proposed text detection and recognition algorithm achieves higher accuracy.

Lovell et al. [4] Instruments that perform fundamental activities such as edge, pattern, and rotation detection can help to streamline a process. To locate dials and segments and estimate their values, circular edge detection, rotation detection, and pattern searches are used. The values of a display are identified and determined using vertical and horizontal edge detection, as well as a region of interest. They demonstrate how to read metre data.

Cerman M. et al. [5] provided a method for reading electricity, gas, and water metres using a mobile platform. The suggested method is divided into two stages: digit detection and optical character recognition (OCR). The input image is subjected to a series of processes in order to detect the digits. Two distinct techniques, Tesseract OCR and CNN, were used in the OCR stage, and they were compared. The accuracy of both the digit detection and OCR stages was found to be high after testing a large number of real-world photos.

III. Proposed System

In this project, the user enters a scanned image as input through an interface. The interface for the proposed system is a website application. This image is then processed using various image preprocessing methods such as background noise removal, gray scaling, histogram processing. To find the required part from the input image, edge detection methods are used. The preprocessed image is then fed to a machine learning model. This machine learning model is highly trained using the convolutional neural network technique. The ML model uses a digit data set for training and testing. The ML model predicts the output as text from the input image. This text is the output of the system and is stored automatically in the database, and the same output is provided to the user.

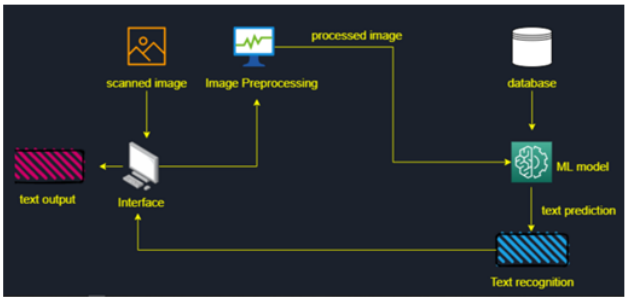

Fig. 1. System Architecture

Our system is separated into three key elements, as illustrated in the diagram:

Image scanning: Image scanning is done by the user. The user needs to use a device camera to take the image of the object to be scanned.

Image preprocessing: This part consists of applying different image preprocessing methodologies to get rid of background noise from the image. The image will be enhanced in order to extract maximum data from it.

Text output: The predicted output will be stored in the database. The same output will be shared with the user through the same interface.

The system contains mainly five phases:

A. *Image preprocessing*

Inputting an original image to the prediction model directly without any processing will not bring very accurate results. To increase the reliability of our model, it is extremely critical to process our image by providing necessary functions like grayscale, thresholding, inversion, and sharpness & blurriness in order to make the image suitable for an OCR model. Applying these functions improves the accuracy of the prediction of the text in the image. This will help us achieve a higher percentage of accuracy of the text prediction model.

Grayscale - Grayscale is the value of each pixel in a digital image that simply represents light intensity information. Only the darkest black to the brightest white are usually displayed in such photos. That is, the image only has black, white, and grey hues, each of which has various levels of grey.

Inversion - Due to physical limitations of the measurement devices, inverse problems involve estimating parameters or data from insufficient observations. The observations are often noisy and contain incomplete information about the target parameter or data. Deconvolution is a technique for extracting clear images from a series of hazy observations.

Thresholding - Image segmentation is an example of thresholding, which involves changing the pixels of an image to make it easier to analyse. Thresholding is the process of converting a colour or grayscale image into a binary image, which is just black and white.

Blurring - The Gaussian blur is a low-pass filter that can be applied to an image. It is used to remove random noise from the image. A median filter is generally employed for different types of noise, such as "salt and pepper" or "static" noise.

Smoothening and sharpening - Preprocessing techniques such as colour image smoothing are used to remove possible image disturbances without sacrificing image information. Sharpening, on the other hand, is a preprocessing technique that aids in feature extraction in image processing.

B. *Edge Detection*

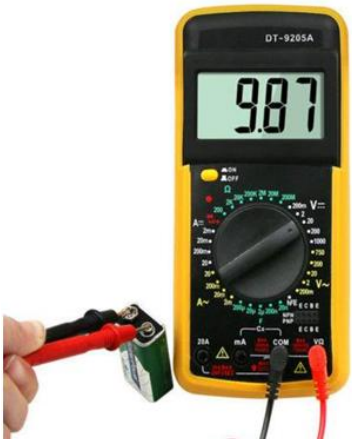

Fig. 2. Edge Detection - Original Image

This technique is used to capture the region of interest ie. the digital display of the scanned device.

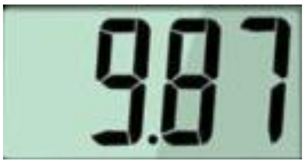

Fig. 3. Edge Detection - Cropped Image

C. Character Segmentation

Segmentation has been done on the region of interest approach. First, the read image is converted into grayscale. Then thresholding and Gaussian blurring is applied so that the characters on the display are revealed as a binary image.

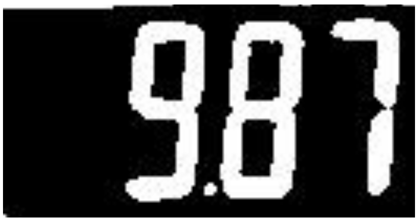

Fig. 4. Character Segmentation - Binary Image

To identify the charachters, the pixel values is seen. Using the technique of connected -component analysis character with large white areas can be determined.

For this every character has a particular range of values. In the connected-component approach the lower boundary and higher boundaries are determined and the character is separated.

Finally, each connected-component is looped-over, and characters are determined.

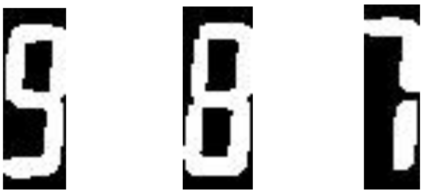

Fig. 5. Character Segmentation - Segmented Characters

D. User Interface

A user interface is a website that connects the user to the intelligent system. The user will be able to scan an image using the interface. The website will give an option to capture an image. The website will require camera access permission in order to use the system camera.

After taking a picture of the meter, there are two options. The user can retake the image if he is not satisfied with the quality of the image captured. If the image is in good shape, the user can choose to transmit it to the intelligent system for data extraction by clicking on proceed.

The machine learning model extracts data from the image and produces text, which is saved in the database and shared with the user.

E. Data extraction using ML

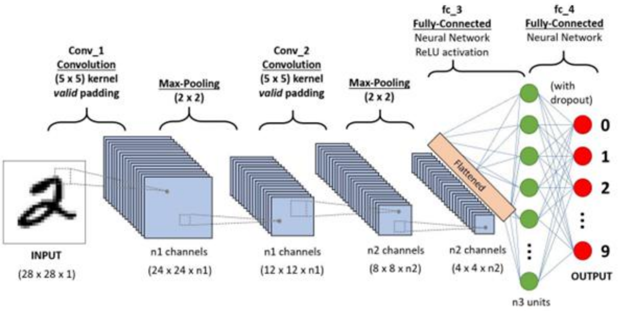

Fig. 6. Convolutional Neural Network

Input layer: The first layer of the model takes a multi-dimensional array of input image. The software uses a matrix filter to perform convolution on the input image.

Max-pooling Layer: The input layer is followed by the max-pooling layer, where the software takes height and width from the input image. If the image is previously recognized by the system, then it simply compressed into less detailed image.

Prediction Layer: This layer compares the output to the input text. The output is fed into forward neural network and the output is predicted using the softmax activation function.

IV. CONCLUSION

Intelligent systems are being used in the industry for data extraction and automating the data handling processes. Automation helps reduce human errors and saves time. Advanced machine learning algorithms are being used to make powerful models that extract and recognize data precisely.

The current project provides a machine learning-based intelligent system that takes a preprocessed image and extracts information from the image. The input image can contain background noise and other issues. Image preprocessing is required for such image inputs before they are fed to the machine learning model. The machine learning model is trained and tested using a character dataset and convolutional neural network. The ML model extracts data. This extracted data is stored automatically in the database and shared with the user.

V. FUTURE WORK

We intend to increase the efficiency of the machine learning model by training the model with different variety of character training images. We also intend to work on decimal readings and recognize alphabetical characters.

ACKNOWLEDGEMENT

This project necessitates a high level of technical skill as well as understanding of image processing and computer vision, and tremendous support of guides, and the success of a project like this is only possible with every individual’s efforts. We take this opportunity to thank everyone who has contributed to the project's success. We would like to express our gratitude to Mr. Ravish Sahay for his unwavering support and assistance. We were able to learn a lot from this project thanks to his prompt instruction and support. Our gratitude and appreciation also go to my project coworkers and Ms. Dipti Jadhav, who has volunteered her time and talents to assist us. We would also like to thank Ms. Priya Kaul, our project coordinator, for giving us with regular updates on documentation and the project timeline. She made certain that we were always on time. A big thanks to the staff at the Don Bosco Institute of technology for allowing us unrestricted access to all the resources available during the course of our project work. Thank you everybody for your assistance in making this possible.

REFERENCES

[1] M. Geetha, R C Pooja, J. Swetha, N. Nivedha, T. Daniya. (2020). Implementation of Text Recognition and Text Extraction on Formatted Bills using Deep Learning. International Journal of Control and Automation, 13(02), 646 - 651. Retrieved from http://sersc.org/journals/index.php/IJCA/article/view/11207..

[2] Mande Shen and Hansheng Lei, "Improving OCR performance with background image elimination," 2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 2015, pp. 1566-1570, doi: https://doi.org/10.1109/FSKD.2015.7382178.

[3] Tran Duc Duan, Duong Anh Duc and Tran Le Hong Du, "Combining Hough transform and contour algorithm for detecting vehicles' license-plates," Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, 2004., 2004, pp. 747-750, https://doi.org/10.1109/ISIMP.2004.1434172.

[4] B. Tejas, D. Omkar, D. Rutuja, K. Prajakta and P. Bhakti, "Number plate recognition and document verification using feature extraction OCR algorithm," 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), 2017, pp. 1317-1320, https://doi.org/10.1109/ICCONS.2017.8250683.

[5] Cerman M., Shalunts G., Albertini D. (2016) A Mobile Recognition System for Analog Energy Meter Scanning. In: Bebis G. et al. (eds) Advances in Visual Computing. ISVC 2016. Lecture Notes in Computer Science, vol 10072. Springer, Cham. https://doi.org/10.1007/978-3-319-50835-1_23.

[6] Karthick Kanagarathinam, Kavaskar Sekar, Text detection and recognition in raw image dataset of seven segment digital energy meter display, Energy Reports, Volume 5, 2019, Pages 842-852, ISSN 2352-4847, https://doi.org/10.1016/j.egyr.2019.07.004.

[7] J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73..

[8] Lovell, T., & Litwhiler, D. (2006, June), Teaching A Computer To Read: Image Analysis Of Electrical Meters Paper presented at 2006 Annual Conference & Exposition, Chicago, Illinois. 10.18260/1-2--324.

[9] Ideal, “A How-To Guide For Using A Recruitment Chatbot”, https://ideal.com/recruitment-chatbot/, 2020 (last accessed: Dec 2020).