Intelligent Entity Information Retrieval System Based on Text Detection and Recognition for Smart Phone Devices

Pushpendra Kumar, Kamal Kant Sharma, Mukul Aggarwal, Varun Kaushik

Abstract*: Text Detection and recognition can be used in several ways to help smart devices. The capability of text detection and recognition can prove an intelligent data input system and can reduce the time of storing any information about any entity to a large extent. Sometimes, it tends to be very boring to input the data about any entity manually. This data can turn into useful information in future. There are several possible entities of interest in real world about which user can be interested. The data related to this chosen entity can be managed and maintained by user easily. User can also export extracted data to different formats according to requirement.

***Index Terms*: text-detection and recognition, machine-learning, Firebase, supervised-learning algorithms**

I. INTRODUCTION

Machine learning [1] is providing several new abilities to simple electronic devices. A smart phone can search google by speech recognition. You don't need to input typing manually. You just need to ask for the phrase or word you want to search. Machine learning enabled device will take care of all the complex computation in background. User can access the information more easily.

We can use this ability of machine learning to save, retrieve and share the information of various entities in much easier way. For this intelligent entity information management system we are using Text detection and recognition. We can use this capability to fetch digital information about any real world entity.

Entity is an object or an event which can be abstract or concrete. An attribute of an entity is important to categorize different entities. An attribute is a feature or property of an entity. The common user can understand the real world more easily with the help of this representation.

For example reading and saving data digitally from a business card can be time consuming, but by using text detection and recognition intelligently we can improve the speed of data input considerably. As a business card being an entity can have several fields. We don't need to fill the values of these fields manually. We can customize our fields and add new fields to given entity and make information comprehensive and useful. We can also export this information to different formats also. There can be several entities like person, business card, place, building, office, relative, friend etc. Users must be able to create their own entities of interest.

The objective of this application is acquiring the data from our environment; this data can turn into some useful information for the user in future.

We also need to recall few things about smart phones before moving forward. Smart phone are based on an operating system specially designed for this type of hardware capable of running applications. Apple's iOS, and BlackBerry's BlackBerry OS, Google's Android OS, HP webOS, and Microsoft's Windows Phone are few examples. In fact, Recently Android becomes operating System installed on largest active devices. These devices have battery, storage and size constraints. These devices are equipped with cameras, Wi-Fi, Bluetooth, NFC and several other useful sensors.

II. II. REQUIREMENTS, PREREQUISITES AND VIEW OF PROBLEM

In this paper, for implementation purpose we have used firebase [2] [3] ML Kit (Machine Learning Kit). This is a service provided by firebase to use machine learning [7] capabilities in Android and IOS Apps.

As our focus is on extracting the data using text detection and recognition, we can use this library to implement machine learning capability. We are also using Android studio to develop an android App. Since android phone have a camera

[14] device installed, we can use a smart phone loaded with high quality camera. This app has the capability to detect and recognize the text using the camera.

Text recognition can automate tedious data entry for credit cards, receipts, and business cards, or help organize photos.

Revised Manuscript Received on December 22, 2018.

Pushpendra Kumar, Computer Science and Engineering, Meerut Institute of Engineering and Technology, Meerut, India

Kamalkant Sharma, Information Technology, KIET Group of Institutions, Ghaziabad, India

Mukul Aggarwal, Information Technology, KIET Group of Institutions, Ghaziabad, India

Varun Kaushik, Computer Science and Engineering, Meerut Institute of Engineering and Technology, Meerut, India

With the help of Cloud-based API like firebase ML Kit, text can be recognized from documents, which you can use to increase accessibility or translate documents. These

Applications can also trace real-world objects, such as by reading the street number, bus number etc.

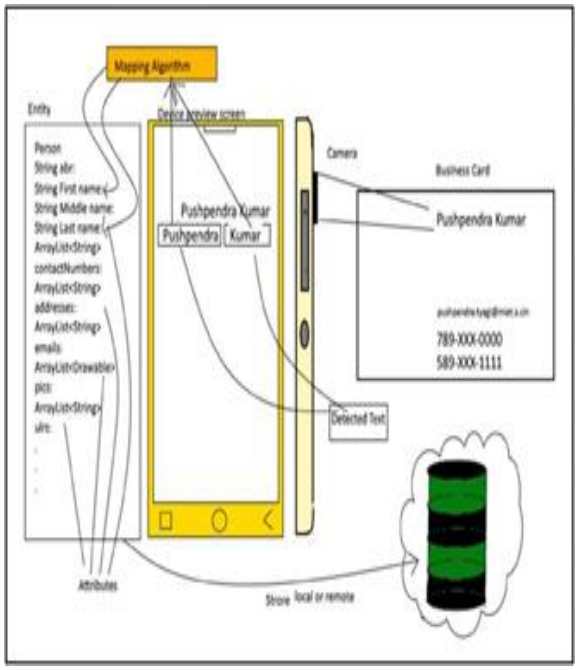

Fig 1. Process of acquiring the data

Detected text can be used to populate the entities attribute fields. Our main concern is to develop an algorithm to provide a correct mapping between the entity attribute and text detected. This algorithm must also have the capability to introduce new attributes to this entity if user wants to add.

There are some applications available for business card scanning or processing, and fetching the data populated over it. This paper presents a much broader and general picture of the problem. Our requirement of data acquisition cannot be limited only to a simple business card. Useful data for the user is available everywhere which can be useful for the user in near future. For example a smart [15] device user who visited a zoo few days back, can be interested in data during his visit. Suppose user saw a lion at zoo. Just clicking a picture cannot obtain complete data most of the time. Now, this application injects the capability of just clicking a snapshot of image having some data like name of the animal, age, size, weight and much more, which was easily available outside the cage of animal. Now user can have much qualified and sophisticated information about lion without doing much work.

There can be several such scenarios. Another simple example is where user is visiting restaurant x and wants to acquire the details of menu, recipes and prices. User can create a new entity called food at restaurant x. User can acquire related data in text format in no time instead of just clicking the snapshots of crockery. The area of data acquisition can be incremented according the requirement of the user.

III. MAPPING ALGORITHM

First user is allowed to describe the name of entity in which user is interested and want to store data about this entity for later use. Now is allowed to move to Live preview screen. This screen is also known as an Activity in Android. Some

Input from live preview is selected from some resource like in our example we have taken the example of business card. However, this resource can be anything like a shine board, advertisement hoarding, label, receipt pamphlet, picture etc. The text can be detected and recognized from this simple scanning process. The text is fetched from this input live preview screen which is designed to do this scanning. Now a mapping algorithm can work to identify the suitable attribute for this text.

The mapping algorithm has the functionality to match the input text using Regex and String matching. Every attribute generally have some characteristic like name of a person does not have a numeric character, email id contains @ character, phone number does not have an alphabetic character etc. These things can be easily validated and verified using Regex.

If a pattern contains a particular Regex than it is a very good candidate for classified attribute of an entity. We have also added few string matching functions to improve the speed and efficiency of this algorithm. For example abbreviation attribute of person. It can be out of limited set of values only like Mr., Mrs., Dr., Prof., etc. There are only few sophisticated abbreviations only, so if a recognized text lies in this set it can be easily identified by string matching. Another example of this matching is contact numbers, as they are generally in fixed length.

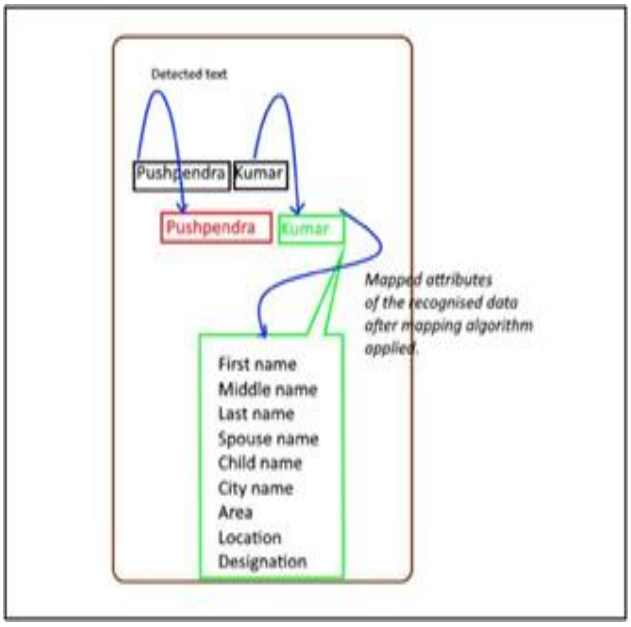

Fig 2. Suggesting attributes on the basis of detected data

Mapping algorithm suggests best suitable attribute value combination to the user. This combination is displayed by another graphical interface to the user. Now user can also intimidate this list. User can edit and modify this best suitable list using drag and drop option. Drag and drop is very user friendly operation for touch screen devices. We can reduce the time of interaction with device to a great extent.

To save still images, Audio/Video we can use the extended facility of platform in which the application is developed.

We can provide three type back end support to this application.

a. Locally (XML format, JSON format, IO files etc.)

b. Locally/ remotely in a database

c. Remotely in cloud base infrastructure [9]

Each of this method has its own advantages and disadvantages, however storing data remotely on a cloud based facility are much more useful in terms of high data availability.

IV. IMPLEMENTATION OF MAPPING ALGORITHM

Implementation part of mapping algorithm can pictorially represented in fig 3. This algorithm can be started with an empty list of suggestions. List is easy to use e.g. we can use inbuilt List from Collection framework in case of java.

We name list as SL, which is empty initially. The recognized text is accepted by the mapping algorithm. This text can be tested by the algorithm at different stages and two basic methods minifySL( ) and maxifySL( ) can be applied on SL. Method maxifySL() is applied only once at first stage which imports list of attributes from two different XML files. First file contains all the attributes defining numeric text and other file contains all the attributes defining non numeric text. For example all contact numbers come in category of numeric text and all email ids come in the category of non-numeric text. Now, SL is populated with lot of predefined entries. We need to minimize the entries and suggest them to the user. At all other stages we minimize SL using minifySL( ) method. This is a sequential algorithm and introduces these stages to algorithm one by one. So we can reduce and add stages to the algorithm. At each stage if the criteria are met then we can minimize the SL and can move to next stage. At final stage we get the minimized list. This list can be displayed to the user on screen. However the algorithm is strong enough to leave user to a single possible attribute but if is not able to do so user can create a new attribute. This new attribute will be now added to one of two XML files and system will be trained with new attribute value.

The conditions tested at each stage are tested using regex or String matching algorithms. Regular expressions or regex are used by search engines, word processors and text editors frequently. They are used to search and replace fragments of words and strings. These are very useful in text processing. Utilities such as sed and AWK are used in lexical analysis. Many programming languages such as Java provide regex capabilities, built-in or via libraries.

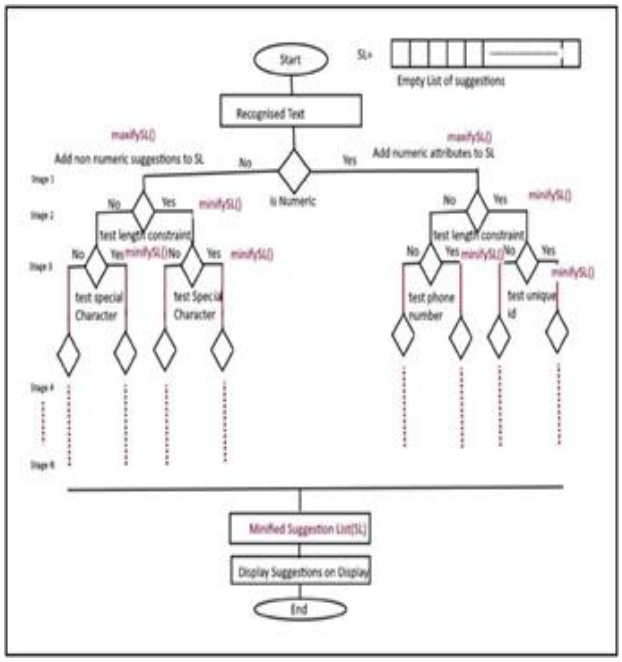

Fig 3. Mapping Algorithm

Mapping algorithm can improve itself using supervised learning [10] .The algorithm generates a function that maps inputs to desired outputs. The results can be used to enhance the dataset available in XML files to generate better suggestions.

VI. CONCLUSION

Nowadays, Digital devices loaded with powerful digital camera, accelerometer, GPS, gyroscope, proximity sensors, high resolution, extremely touch sensitive screens and so many other advanced hardware components along with extremely agile processors and fantastic operating systems are available with reasonably low prices. The capability of these devices can be increased to great extent by using the features like machine learning. This paper presents a bridge algorithm to extract the functionality of two powerful technologies, smart digital devices and machine learning. We introduce a decision making capability to an application powered with machine learning with some simple methods like pattern recognition and string matching.

V.FUTURE SCOPE

This paper is more about the idea of how to store useful information in convenient manner for the user. We want to reduce the burden and length of process of saving the information.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

This process can get better with both hardware and software supplements involved in this data acquisition process. If quality of camera [14] improves the text detection and recognition will get faster and better. Also, if the mapping algorithm improves the number of suggestions occurring in front of user will also reduce. The selection will become easier and process will improve significantly.

However, mapping algorithm performance depends on the dataset which it has learned and is populated previously; still minimized dataset is good for the speed and efficiency. So, there is future scope for another supporting algorithm which can introduce the capability to keep this dataset in check.

REFERENCES

1. Alex Smola and S.V.N. Viswanathan, "Introduction to Machine Learning" in the press of the syndicate of the university of Cambridge,ISBN 0 521 82583 0, 2008.
2. http://lp.google-mkto.com/rs/248-TPC-'286/images/Google-Next-2007-Keybook.pdf
3. http://firebase.google.com/docs/ml-kit/recognize-text
4. http://en.wikipedia.org/wiki/Regular_expression
5. https://msdn.microsoft.com/en-us/library/twcw2f1c(v=vs.110).aspx
6. Qixiang Ye,David Doermann, "Text Detection and Recognition in Imagery: A Survey" IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014.
7. J.J. Weinman, ZButler, D. Knoll, J. Field, " toward Integrated Scene text Reading", IEEE Transaction Pattern Analysis and Machine Intelligence, Vol-3, No-2, pp. 375-387, 2014.
8. S. Wachenfeld, H. Klein and X. Jiang, " Recognition of Screen-Rendered Text" , in proceedings IEEE International conference on Pattern Recognition, pp. 1086-1088, 2006.1733-1746, 2009.
9. R. Buyya, C.Yeo, S. Venugipal, J. Broberg & I. Brandic, "Cloud computing and Emerging IT Platforms: Vision, Hype and reality for delivering computing as the 5th Utility, Future Generation Computer Systems", 25(6):599-616, Elsevier, The Netherlands, 2009.
10. Taiwo Oldipupo Ayodele,"Types of Machine Learning Algorithms", University of Portsmouth United Kingdom, 2010.
11. T. M. Mitchell, "The Discipline of Machine Learning Department technical report", Carnegie Mellon University, 2006.
12. V.J. Hodge and Jim Austin, "A survey of Outlier Detection Methodologies", Kluwer Academic Publishers, 2004.
13. http://www.cs.unc.edu/~stotts/145/CRC/eer1.html
14. C. Fiorini, A. Longoni, F. Perotti, C. Labanti, E. Rossi, P. Lechner, H. Soltau, L.Struder, "Silicon Drift Detector Readout Electronics for a Compton Camera", Elsevier, 2018.
15. https://en.wikipedia.org/wiki/Smartphone