World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:13, No:7, 2019

Non-Invasive Data Extraction from Machine Display Units Using Video Analytics

Ravneet Kaur, Joydeep Acharya, Sudhanshu Gaur

Abstract—Artificial Intelligence (AI) has the potential to transform manufacturing by improving shop floor processes such as production, maintenance and quality. However, industrial datasets are notoriously difficult to extract in a real-time, streaming fashion thus, negating potential AI benefits. The main example is some specialized industrial controllers that are operated by custom software which complicates the process of connecting them to an Information Technology (IT) based data acquisition network. Security concerns may also limit direct physical access to these controllers for data acquisition. To connect the Operational Technology (OT) data stored in these controllers to an AI application in a secure, reliable and available way, we propose a novel Industrial IoT (IIoT) solution in this paper. In this solution, we demonstrate how video cameras can be installed in a factory shop floor to continuously obtain images of the controller HMIs. We propose image pre-processing to segment the HMI into regions of streaming data and regions of fixed meta-data. We then evaluate the performance of multiple Optical Character Recognition (OCR) technologies such as Tesseract and Google vision to recognize the streaming data and test it for typical factory HMIs and realistic lighting conditions. Finally, we use the meta-data to match the OCR output with the temporal, domain-dependent context of the data to improve the accuracy of the output. Our IIoT solution enables reliable and efficient data extraction which will improve the performance of subsequent AI applications.

Keywords—Human machine interface, industrial internet of things, internet of things, optical character recognition, video analytic.

I. Introduction

THE advent of Internet of Things (IoT) has seen fundamental changes in data acquisition, storage and analysis to create new value in multiple verticals such as manufacturing, energy, oil, gas, transportation and so on. IoT in conjunction with machine learning and AI has the potential of transforming society and improving human life. Though the term IoT has become a ubiquitous buzzword in recent times, it encompasses many different aspects that depend on the specific use case and technology. The challenges and constraints of a successful IoT deployment is also not unique but variable. A useful (though still broad) categorization is to contrast between IoT solutions for new greenfield applications (such as smart cities, smart home, autonomous driving) versus legacy brownfield applications (such as manufacturing, oil and gas). The latter is often termed as Industrial IoT (IIoT) [1] to distinguish from the former and typically relate to critical societal infrastructure deployments which have to be in continuous 24×7 operation. Any introduction of IoT for these use cases have to respect existing reliability, availability and safety constraints before attempting to improve aspects of existing operations [2].

Ravneet Kaur is with the Research and Development Division, Hitachi America Ltd., Santa Clara, CA 95054 USA (phone: 408-859-3792; e-mail: Ravneet.Kaur@hal.hitachi.com).

Joydeep Acharya and Sudhanshu Gaur are with Research and Development Division, Hitachi America Ltd., Santa Clara, CA 95054 USA (e-mail: Joydeep.Acharya@hal.hitachi.com, sudhanshu.gaur@hal.hitachi.com).

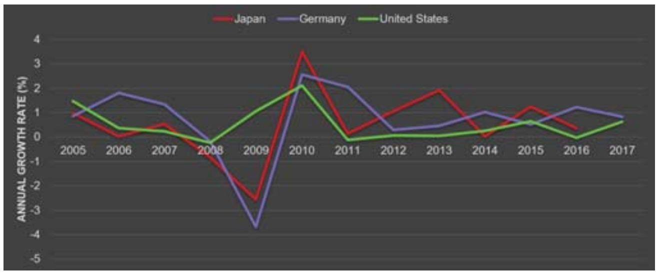


Fig. 1 Manufacturing multi-factor productivity trends for three markets (Figure reproduced from [3])

For many such legacy societal infrastructure deployments, which is also called Operational Technology (OT), data acquisition and storage in an Information Technology (IT) database can pose a huge challenge. This could be due to legacy OT protocols for data acquisition which were designed for in-situ data visualization and manual recording. Security concerns may also limit direct physical access and networking to these controllers for data acquisition. However, having access to data in an IT database is the pre-requisite for building a successful AI solution. In this paper, we have addressed this issue and proposed an IIoT technology for data extraction and storage that does not need physical connection to the existing OT network or affect its operations in any way. Hence, we term this as a *non-invasive data extraction method*. This ensures that the reliability, availability and safety constraints of the OT networks are respected while, the OT data are connected to an AI engine.

II. Manufacturing Landscape

To understand our proposed solution in more concrete terms, we shall consider its applications and value in smart manufacturing [4]. As seen in Fig. 1, the multi-factor productivity in manufacturing (which reflects the overall efficiency with which labor and capital inputs are used together in the production process) has been stagnant for the last decade. A main reason is that the various operations of the shop floor (also called plant or production floor) have reached a ceiling in terms of further optimization. Such operations encompass production, maintenance and quality related tasks. The next wave in terms of optimization will come from operationalizing AI and analytics for shop floor operations. A primary bottleneck is the lack of availability of high quality data with sufficient volume, velocity and variety.

A. Data in Manufacturing

The basic source of data in manufacturing comes from the industrial controllers such as Programmable Logic Controllers (PLC) and/or Supervisory Control and Data Acquisition (SCADA) systems. In factories, there are a multitude of machines that perform important manufacturing processes. The actions of each machine (or group of machines) is controlled by a PLC that issues corresponding commands. Such actions could be to turn a machine 'ON' or 'OFF', increase linear velocity of a belt, rotational velocity of a turbine, controlling parameters for picking and placing actions of a robotic arm and so on. A PLC also captures sensor information that indicates machine states such as temperature, pressure, count of parts moving along an assembly line, and so on. All these values (sensor and command) are written in PLC memory units. A machine command would in turn effect the sensor readings. For example, a sensor would measure non-zero belt velocity after the command to turn it ON.

How can the data about the sensor and commands be acquired and connected to a subsequent AI application? In principle, this is where a SCADA system is used. A SCADA system is a plant-wide software system that can be used to program the PLCs and to acquire the data that the PLCs obtain from the machines. The data for a subsequent AI application can be acquired by querying the SCADA database. However, as we shall see in the next subsection, this may not always be straightforward.

B. Challenges in Data Acquisition

There are multiple challenges in seamlessly connecting OT data to an AI application. Some of these are listed here,

1) In many factories, the entire PLC data is not being currently stored in a SCADA system. The reason is that the current operations rely on manual data acquisition by reading data from the display units attached to the PLCs. This practice also contributes to the stagnant multi-factor productivity which was mentioned above.
2) Even in factories where plant managers realizes that manual data acquisition isn't optimal, there could be lack of time and resources to reprogram PLCs plant-wide to make them capable of automatic data acquisition in a SCADA system. Different PLCs can have different programming environments with varying degrees of complexity and feasibility. Some machines can be operated by very special custom controllers and reprogramming them would require contacting the manufacturer.
3) Re-programming PLCs for purpose of data acquisition will require stopping and restarting the PLCs which will affect the 24×7 availability and reliability of the factory operations, something which many plant managers are loathe to agree upon.
4) Even if all PLCs are re-programmed to store all their data in a SCADA system, a plant manager would rather use that data for daily factory operations. The manager could still be wary of connecting this data to an IT system where, the AI module will reside. The reason for this is security and the risk of infecting the OT network with virus such as Stuxnet.

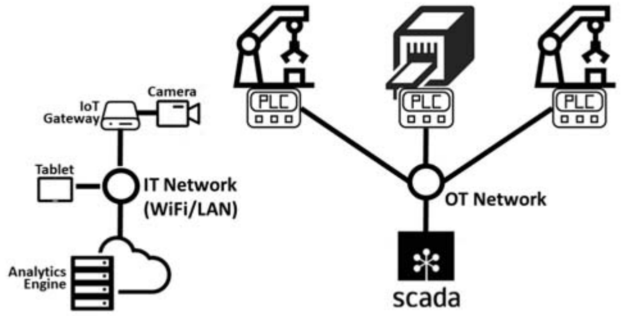


Fig. 2 High-level overview of non-invasive data extraction

III. Non-invasive Data Extraction

To address the data acquisition challenges mentioned in the last section, we propose an IIoT solution which positions a video camera in front of the digital display units attached to the PLCs and runs an edge analytics algorithm at a connected IoT gateway to automatically extract and store the values being displayed on the screen. The high-level overview of this solution is shown in Fig. 2. Note that the camera, gateway and analytics engine are connected to an IT network which is separate from the OT network which connects PLCs and SCADA server. Since, there is no need to connect the two networks, we call our solution as non-invasive.

Our solution uses Optical Character Recognition (OCR) to interpret the digits on the screen. However, simply using an OCR algorithm doesn't guarantee a reliable solution due to the complexity and variability in digital display units found in factories. Our solution augments the OCR algorithm with additional modules as shown in Fig. 3.

The function of three modules are briefly mentioned here and are explained in detail in section IV.

1) Image Pre-Processing: We provide the appropriate stakeholder in the factory with a tablet running an android application. The application allows its user to view the list of available machines and select one for which non-invasive data extraction is desired. A static snapshot of the controller screen for that machine appears on the application and the user can annotate the Regions of Interest (RoI) in the screen where the time-varying data is located. The coordinates of the RoI is sent to an edge analytics module that resides in an IoT gateway. This is a one-time operation. For our implementation, we assumed a MQTT over WiFi connection.

2) Optical Character Recognition: The edge analytics module in the IoT gateway runs an OCR algorithm to the RoI. Thus, the OCR doesn't have to be performed on the whole complex display screen and the performance is enhanced.

3) Data Post-Processing: The edge analytics module also calculates the confidence level [5] of the OCR result. This is stored in the IoT gateway along with the OCR result and is used as an additional input to a subsequent analytics application that uses the data.

Fig. 3 Functional blocks in proposed non-invasive data extraction solution

IV. Detailed Explanation of Solution

In this section, we provide details of the steps mentioned in the previous section.

A. Image Pre-Processing

The screenshots of the mobile application is shown in Fig. 4 where `Machine List` consist of all the machines and their respective PLCs (we skip the details of how to populate this list in this paper for sake of brevity). The user of the application can select a machine for which non-invasive data extraction is desired. For example, in Fig. 4a, a PLC called XCRON is selected which then pulls up a static snapshot of the Human Machine Interface (HMI) display on the application. The HMI shows various information stored in the XCRON PLC counters. A user can now annotate the RoI based on the specific requirement about which of these counter information need to be acquired. To do this, the application provides a `Select Area` button which pops an annotation box (in red as shown in Fig. 4b) for annotating the RoI by dragging it on top of the RoI. The annotation box can be formatted to a specific size to provide more precise location of the RoI.

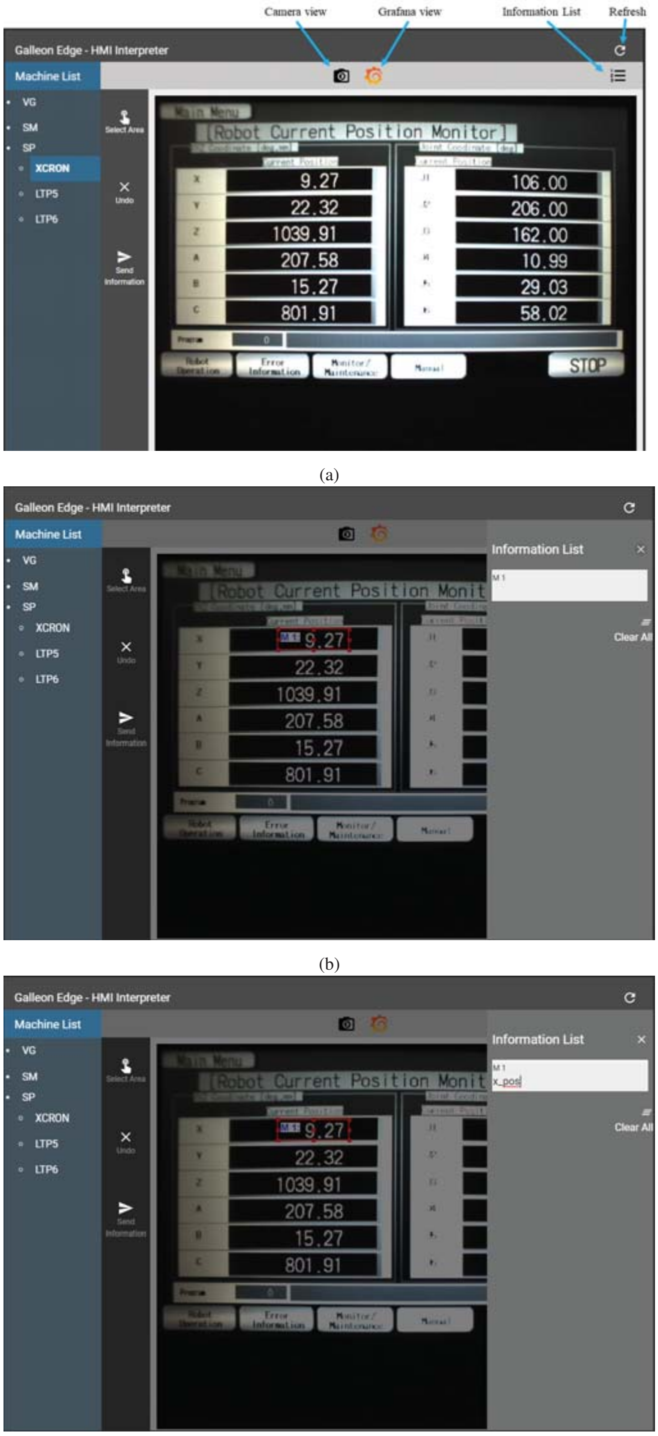
After user has selected the desired RoI, an information box is prompted where user can specify a label for each of the selected regions (e.g. x_pos in Fig. 4c). The user can also modify the labels at a later time. After annotating the RoI, the user can then press the `Send Information` button to sent the coordinates of the annotated region and the tag/label to the edge analytics algorithm.

B. Optical Character Recognition

Optical Character Recognition (OCR) digitizes an image text for computer manipulations. Tesseract [6] and Google Vision [7] are two OCR engines with high performance for a wide range of applications. In this paper, we have implemented and compared tesseract and google vision for the images extracted from the camera attached in front of the HMI.

1) Tesseract: Tesseract [6] is an open source OCR engine that has the capability to recognize the traditional black on white characters as well as white on black text. The back-end of the current architecture for tesseract is as follows:

- The captured image is first converted to gray-scale and the regions were cropped from the image using the coordinates defined by the user.
- Thereafter, Otsu thresholding [8] is applied on the cropped image which contains the desired text.
- At times, the segmented images may contain borders that may interfere with the recognition. Hence, these borders were removed by first finding the contours and then recognizing the contours that are close to the borders and finally removing the recognized borders from the image [9].



(a)

(b)

(c)

Fig. 4 Screenshots of the mobile application with (a) an image of a PLC, (b) an annotated region with a fixed tag, (c) a user-defined label for the annotated region

Fig. 5 Screenshot of the mobile application with graphical representation using Grafana

- Tesseract is then applied on the cropped-segmented-border-removed images to recognize the characters.

2) Google Vision: Google vision [7] is an application programming interface (API) that supports text detection (extracts text from images) and document-text detection (extract dense text from an image). The back-end of the current architecture for google vision is as follows:

- The captured image is first converted to gray-scale and the regions were cropped from the image using the coordinates defined by the user.
- Thereafter, Otsu thresholding [8] is applied on the cropped image which contains the desired text.
- Google vision is then applied on the cropped-segmented images to recognize the characters.

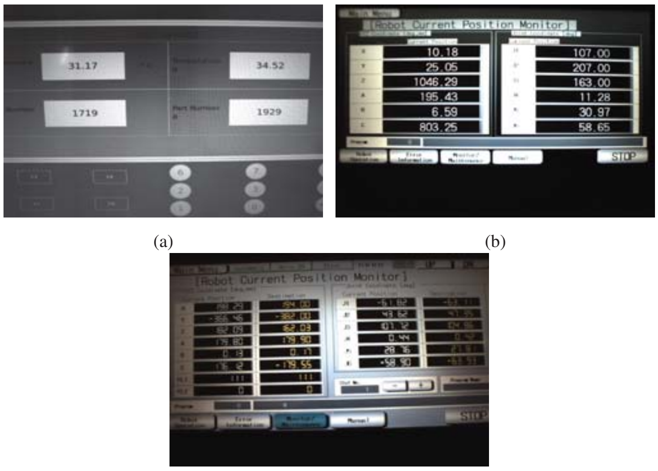
In this research, our aim is to provide the complete solution that will work for a wide range of manufacturing environments and not necessarily optimize the OCR algorithm performance for a given environment.

C. Data Post-Processing

In manufacturing facilities, the data generated by the PLCs are mainly in the numerical form. Hence, we have post-processed the recognized characters to be identified as numerical characters. We have performed a numerical character verification to ensure that each character in the annotated region is a number and if, any recognized character in a cropped-image is an alphabet/symbol then the recognized characters for that cropped-image will not be displayed/passed to Grafana [10] for viewing. Apart from numerical verification, confidence level of each character is also taken into account.

In OCR, the degree of characters being correctly recognized is indicated by the confidence rate. In tesseract, a built-in function is used to calculate the confidence of each annotated region. However, an average of confidences for annotated regions in an image is calculated and further discussed in section V.

After the numerical character verification, the recognized characters for an image is stored in the edge analytics module. The mobile application in the tablet can access the stored data for visualization (see Fig. 5). The edge analytics module also makes the stored data available to a subsequent AI application.



(a) (b) (c)

Fig. 6 Camera views of the three PLCs: (a) HMI-1 with 4 RoIs and Roboto font (b) HMI-2 with 15 RoIs and Gothic font, (c) HMI-3 with multiple RoIs and 7-segment (digital) font

V. Results and Discussion

We have evaluated our proposed solution for PLC display units that are used in manufacturing with realistic environment conditions. An example of the images collected from three PLC display units are shown in Fig. 6. The first HMI (HMI-1) contains 4 RoIs with Roboto font, the second HMI (HMI-2) contains 12 RoIs with Gothic font, and the third HMI (HMI-3) contains multiple RoIs with 7-segment (digital) font.

The static snapshot of these camera images could be accessed through the mobile application for the user to annotate the regions of interest. Subsequently, images were collected every second from each of the three PLCs and the OCR accuracy was calculated. The two OCR engines, Tesseract and Google Vision were implemented and compared for the images collected from the camera of the three PLCs.

For HMI-1, all 4 RoIs were extracted for 100 seconds; thus for the 100 images, 400 RoIs were extracted from HMI-1. For uniformity in comparison, only 4 RoIs were also extracted from each image for HMI-2 and HMI-3, respectively. The accuracy of tesseract and google vision for the three HMIs were calculated by taking a ratio of the number of correctly recognized characters to the total number of actual characters for 400 RoIs. The accuracy comparison between tesseract and google vision is shown in Fig. 7. For google vision, the images were pre-processed to apply only Otsu thresholding and no other alterations were made to the images. With minimum pre-processing, google vision was able to achieve 97.43% and 99.50% for HMI-1 and HMI-2, respectively. However, HMI-3 which has a 7-segment font, tesseract was able to achieve more accuracy for recognizing characters than google vision.

Seven-segment font posed a challenge. Initially, tesseract was unable to recognize the 7-segment font and the accuracy was about 12%. However, by training the tesseract model for 7-segment font, accuracy was increased to 55.20%. No additional training was performed for google vision.

Note that the performance of tesseract can be further improved by fine tuning the parameters for a given display.

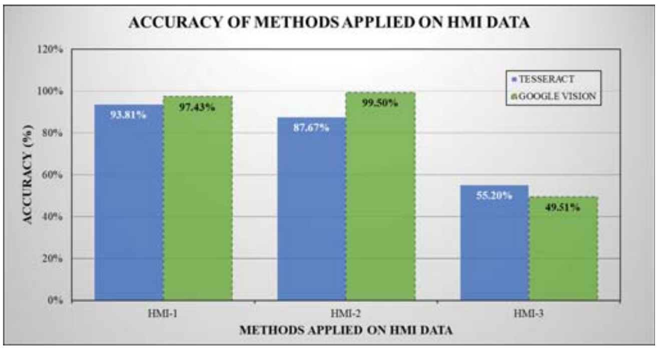


Fig. 7 Accuracy of different HMIs for Tesseract and Google Vision

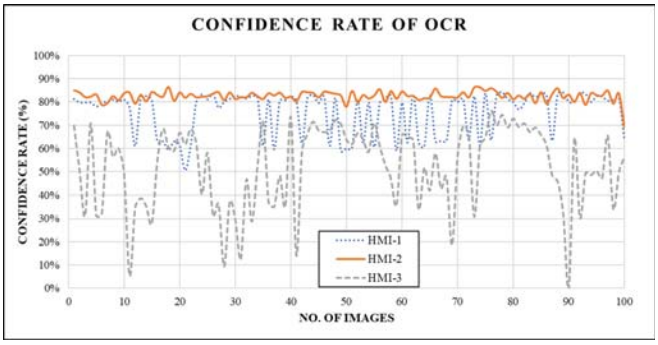


Fig. 8 Average Confidence of OCR for different HMIs using Tesseract

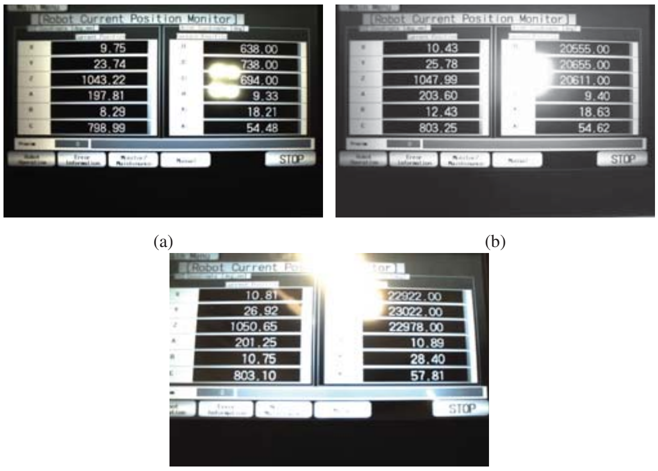
However, our aim in this study is not to find the most optimal setting for tesseract for a given display (each display may require specific configuration) but, to present the solution framework that is applicable in general. Thus, we have moderately trained tesseract to ensure that it can work with the wide range of displays. We can deeply train tesseract for a given display but, then our solution will not be scalable to other displays.

Similar to calculating accuracy, confidence levels for tesseract was also taken into account for the three HMIs. For 100 images, 4 RoIs were extracted from each image of HMI-1 and similarly, only 4 RoIs for each image were also extracted from HMI-2 and HMI-3, respectively. The average of confidence rate for 4 RoIs for each of the 100 images is shown in Fig. 8. HMI-1 and HMI-3 have many fluctuations in confidences and are spread between 0% to 70-80%. However, for HMI-2, the confidence level is within 75% to 85% which shows that for different PLC display units different pre-processing steps are required or even training of the tesseract model is necessary to achieve high confidence for every RoI.

In real world scenario, there may be many environmental conditions that will alter the results and are required to be taken into consideration for optimizing the OCR performance.

1) Optimizing OCR Performance in Factories: In factories, the PLC display unit may not be placed at an optimum location to ensure that there are no interruptions/artifacts that may cause a decrease in the accuracy/confidence of the characters being recognized. One of the known interruptions is the environment light (for e.g. sunlight, ceiling light, etc.). The PLC displays are equipped with an internal back light which is sufficient for the cameras to provide images with sharp character recognition. However, these PLC displays are not equipped with anti-glare/polarized screens and the environmental lighting conditions may reflect the light back to the camera which may result in overexposed images. We have experimented with three different light sources that are being directed onto the HMI-2 and their respective overexposed images are shown in Fig. 9.

Tesseract and Google Vision is also applied on the overexposed images with the same pre-processing steps, and



(a) (b)

(c)

Fig. 9 Overexposed images of HMI-2: (a) Light source 1, (b) Light source 2, (c) Light source 3.

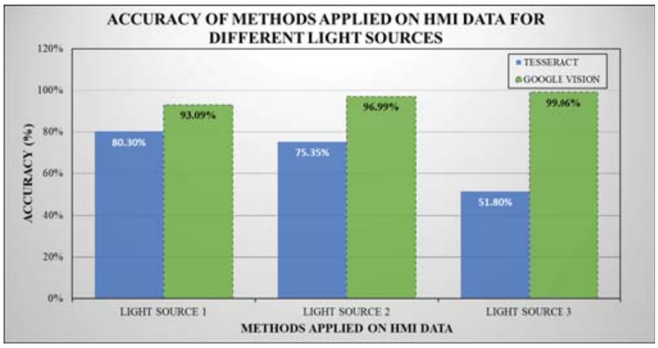


Fig. 10 Accuracy of HMI-2 with different light sources for Tesseract and Google Vision

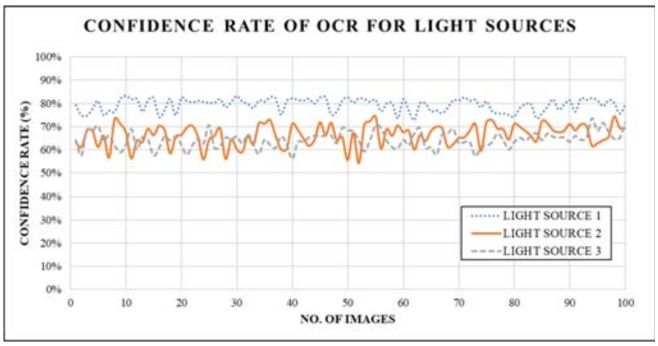


Fig. 11 Average confidence of OCR for HMI-2 with different light sources

the accuracy is calculated and shown in the Fig. 10 for each of the three types of the overexposed images. A total of 100 images were extracted for each of the light sources and passed through the mobile application for user to annotate the RoIs. Here, all 12 RoIs were annotated and sent for character recognition to tesseract and google vision. The accuracy for tesseract and google vision were calculated and google vision yielded 93.09%, 96.99% and 99.06% accuracy for the three light sources, respectively. Similarly, confidence rate for the tesseract is also plotted which shows that the confidence of the tesseract varies approx. 10% and confidence rate for each light source is shown in Fig. 11.

Thus, tesseract being an open source software is a most cost effective OCR engine but, possibly less on performance. On a contrary, google vision is a google based API and performs better than tesseract but, can be very costly.

VI. Conclusion

This paper proposes a solution to automatically extract information (meta-data and variable time series data values) from digital displays found in many environments such as factories. The extraction is non-invasive and thus, is very easy to deploy and possess no security risks associated with traditional methods of data extraction which require connection to the machine. We have tested our solution with realistic PLC display units used in manufacturing, under various real-world environment conditions and obtained reasonably good performance.

References

[1] A. Gilchrist, *Industry 4.0: the industrial internet of things*. Apress, 2016, June 28.

[2] J. Wan, S. Tang, Z. Shu, D. Li, S. Wang, M. Imran, and A.V. Vasilakos, *Software-defined industrial internet of things in the context of industry 4.0.*, IEEE Sensors Journal, vol 16, no. 20, 7373-7380, 2016.

[3] [Online]. Available: https://data.oecd.org/lprdty/multifactor-productivity.htm.

[4] R. Burke, A. Mussomeli, S. Laaper, M. Hartigan, B. Sniderman, *The Smart Factory: Responsive, Adaptive, Connected Manufacturing*, Deloitte, 2017. [Online]. Available: https://www2.deloitte.com/insights/us/en/focus/industry-4-0/smart-factory-connected-manufacturing.html

[5] V. Ohlsson, *Optical Character and Symbol Recognition using Tesseract*, Dissertation, Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, 2016.

[6] R. Smith, *An Overview of the Tesseract OCR Engine*, Ninth International Conference on Document Analysis and Recognition (ICDAR 2007), Parana, 2007, pp. 629-633.

[7] [Online]. Available: https://cloud.google.com/vision/docs/ocr

[8] N. Otsu, *A threshold selection method from gray-level histograms*, IEEE transactions on systems, man, and cybernetics, vol. 9, no. 1, pp.62-66, 1979.

[9] A. Dengel, R. Hoch, F. Hones, T. Jager, M. Malburg and A. Weigel,*CH:Techniques for Improving OCR Results* , *Handbook of Character Recognition and Document Image Analysis*, World Scientific Publishing Company, 2007, pp. 227-258

[10] [Online]. Available: https://grafana.com/