

Research Article

An Offline Image Auditing System for Legacy Meter Reading Systems in Developing Countries: A Machine Learning Approach

Natasha Nigar ,¹ **Hafiz Muhammad Faisal** ,² **Muhammad Kashif Shahzad**,³ **Shahid Islam** ,¹ and **Olukayode Oki** ,⁴

¹*Department of Computer Science (RCET), University of Engineering and Technology, Lahore, Pakistan*

²*School of Engineering and Design, Technical University of Munich, Munich, Germany*

³*Power Information Technology Company (PITC), Ministry of Energy, Power Division, Government of Pakistan, Lahore, Pakistan*

⁴*Department of Information Technology, Walter Sisulu University, Mthatha, South Africa*

Correspondence should be addressed to Natasha Nigar; natasha@uet.edu.pk

Received 29 September 2022; Revised 11 November 2022; Accepted 14 November 2022; Published 25 November 2022

Academic Editor: Raid Al-Nima

Copyright © 2022 Natasha Nigar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The developing countries are challenged with overbilling and underbilling, due to manual meter reading, which results in consumer dissatisfaction and loss of revenue. The existing automated meter reading (AMR) solutions are expensive; hence, sample-based manual snap auditing systems are introduced to control such meter reading inaccuracies. In these systems, the meter reader, besides reading, also collects meter images, which are used to manually audit the meter's accuracy. Although such systems are inexpensive, they are limited in their ability to be sustainable and ensure 100% accurate meter readings. In this paper, a novel offline optical character recognition (OCR) system-based Snap Audit system is proposed and tested for its efficient and real-time 100% accurate meter reading capabilities. The experimental results on 5,000 real-world instances show that the proposed approach processes an image in 0.05 seconds with 94% accuracy. Moreover, the developed approach is evaluated with four state-of-the-art algorithms: region convolution neural network (RCNN), nanonets, Fast-OCR, and PyTesseract. The results provide evidence that our new system design along with novel approach is more robust and efficient as compared to existing algorithms by 43.6%.

1. Introduction

In general, employees from utility companies (electricity, gas, and water) record the consumption manually by walking from house to house and/or building to building on monthly basis. The reading is manually entered in the mobile-based meter reading systems along with meter dial pictures [1–3] for later verification and audit purposes due to extended likelihood of error [4–6]. The manual audit of meter readings collected by the field staff on a monthly basis is an expensive and monotonous task in terms of human effort and time [7] and has low accuracy. The utility companies adopt stratified sampling to make a general conclusion about the population, which is based on monthly meter readings. Moreover, manual evaluation of a large number of images may result in negligence of errors.

Besides all benefits and extended analysis support offered by AMR meters [8], upfront cost, infrastructure cost, and technical capacity to sustainable operation are limiting factors for its rollout in the developing countries and the meter reading process is manual. Hence, there is a need to provide a technology solution that provides automation in the existing system without any financial burden or advanced training to operate and sustain while ensuring 100% accurate meter reading. Over the last decade, the trend to digitize paper-based documents has emerged [9]. The aim is to make these documents fully searchable, accessible, and processable in digital form. In this regard, OCR technology has gathered researchers' attention [10]. OCR is the mechanical or electronic translation of scanned images of handwritten, typewritten, or printed text into machine-encoded text [11]. The OCR ideally returns the same output

as input in the digital image. It eliminates inefficient, slow, and error-prone manual processes [12, 13].

The existing approaches for intelligent metering include OCR evaluated on images captured in well-controlled situations [1, 14, 15]. These methods are easily affected by noise and might not be robust to different types of meters (dial or digital with diverse fonts). Furthermore, these approaches cannot solve the problems of skew, dirt, distortion, low light, and uneven light illumination, resulting in low accuracy of the meter reading. Therefore, meter reading recognition under uncontrolled scenarios, e.g., rain, is very challenging and requires employing the latest technology to improve the recognition accuracy. Automatic inspection of images captured by using a meter would reduce the mistakes introduced by the human factor [16] and save manpower [17]. Moreover, the OCR-driven handheld devices are also limited by battery capacity, particularly, when enabled with real-time meter reading extraction from meter dials.

Due to the quality of the meter reading images captured in a complex environment and the error-prone patterns, there is a need to develop an offline OCR that could translate the meter reading images accurately and, hence, reduce labour costs and time. Considering the fact that in developing countries, it is not feasible to quickly replace old meters with smart ones [18, 19]. Therefore, in this study, we present a new system design and a novel offline OCR based on machine learning that processes images captured by a human to replace the manual auditing process carried out by another service company employee. The developed approach utilizes the “You Once Look Only” version 5 (YOLOv5) algorithm [20], which can be applied to various types and models of meters. Moreover, we specifically consider a real-world dataset of 5,000 images. Our developed approach achieves state-of-the-art results by outperforming four existing state-of-the-art algorithms on a real-world dataset. In summary, this work provides the following four main contributions:

- (i) System design: we design a new system that replaces the manual auditing process carried out by a service company employee, requiring human effort and time.
- (ii) Novel offline OCR: a novel approach based on machine learning is developed that automatically recognises the captured meter reading image in an accurate and precise manner, reducing human effort and time.
- (iii) Study of state-of-the-art algorithms: four state-of-the-art algorithms, namely, region convolutional neural network (RCNN) [21], nanonets [22], Fast-OCR [23], and PyTesseract [24] are compared with the developed approach for meter reading recognition.
- (iv) DataSet: the developed approach is tested with 5,000 data images acquired from real-world scenarios by the service company’s employees themselves.

The rest of the paper is structured into five sections. Section 2 overviews the background. In Section 3, we present

the existing system structure followed by the new system design. Experimental analysis and results for meter reading recognition are presented in Section 4. In Section 5, we describe the threats to validity of this study. Finally, Section 6 concludes with possible future directions.

2. Background

The development of machine learning and advances in hardware to process large amounts of image data have influenced researchers for text recognition in digital documents [25]. Meter reading recognition is one of those problems for which machine learning algorithms have been applied [1, 7, 26, 27]. Existing studies analyse the performance of automatic reading systems [1, 14, 15], especially in controlled environments. However, many uncertain events can occur while capturing the meter reading, e.g., rain [28, 29]. In this regard, we investigate the performance of four existing state-of-the-art algorithms to translate the captured image in a complex environment compared to our developed approach. The reason for selecting these four algorithms is that they are simple to implement and most of them are open source.

The RCNN is one of the first large and successful applications of a convolutional neural network (CNN) [30] for solving the problem of object localization, detection, and segmentation [21]. The approach was evaluated on benchmark datasets, achieving state-of-the-art results. In the first step, it extracts region proposals (~ 2000) in different sizes and aspect ratios from the input image. Region proposal is a method where selective search [31] is used to extract the regions of interest (ROIs) from the image [21], which are then fed into CNN to extract features. Feature extraction is a process of dimensionality reduction by which an initial set of raw data is reduced to more manageable groups for processing [32]. In this way, the RCNN yields better results than applying machine learning algorithms directly to the raw data. Next, these extracted features are used to predict the class and bounding box of the region proposal. This process is also illustrated in Figure 1. Due to the classification of 2000 region proposals per image, it takes a huge amount of time (47 seconds for each test image) to train the network, which prevents its implementation in real time.

Nanonets [22] is a powerful OCR tool that uses artificial intelligence (AI) to provide human-level intelligence for data extraction. It leverages deep learning techniques to overcome common data constraints that greatly affect text recognition and information extraction [22]. Therefore, it can recognize images attributed with (i) low resolution, (ii) varying sizes, (iii) shadowy, tilted, or random unstructured text, (iv) noisy, blurred, or images of text in multiple languages at once, and (v) handwritten text. Fast-OCR [23] is a lightweight detection network that incorporates features from existing models focused on the speed/accuracy trade-off, such as YOLOv2, CR-NET, and Fast-YOLOv4 [23]. PyTesseract [24] is an open-source OCR tool for Python that is used to scan and transcribe any textual data in images.

YOLO [33] is a state-of-the-art real-time object detection algorithm. YOLOv5 [20] is an enhancement of



FIGURE 1: Structure of the RCNN model.

YOLOv1 to YOLOv4. The models in YOLOv5 prove to be significantly smaller, faster to train, and more accessible to be used in real-world applications [34]. There are three reasons why we chose YOLOv5 as our first learner. Firstly, YOLOv5 is a merger of the cross-stage partial network (CSPNet) [35] and Darknet, creating CSPDarknet as its backbone. It ensures the inference speed and accuracy along with reducing the size of model. In a meter reading system where millions of consumers' records need to be verified, detection speed and accuracy are imperatives. Secondly, the YOLOv5 utilizes the path aggregation network (PANet) [36] as its neck to boost information flow. PANet helps enhance the location accuracy of the object. Thirdly, the head of YOLOv5, also known as the YOLO layer, enables the model to handle small, medium, and large objects by generating 3 different sizes (18×18 , 36×36 , 72×72) of feature maps to achieve multiscale prediction [20]. The default structure of YOLOv5 is presented in Figure 2. The data are first fed to CSPDarknet, that performs the function of feature extraction, and then input to PANet for feature fusion. Finally, the head, also known as the YOLO layer, detects and outputs the results, i.e., class, score, location, and size.

3. System Design

3.1. Existing System Structure. Handling millions of consumers' records for the billing system is a difficult task. It consumes a lot of human effort, increases costs, and utilizes resources. Figure 3 shows how the existing system works and how the meter reading is currently being taken and verified manually through image auditing. This manual verification is needed to avoid inaccurate readings. The field operation staff captures the meter readings and images on a monthly basis using a mobile application. The captured images are placed on a centralized server. These images are loaded in the desktop application for image-based auditing against the meter readings uploaded manually through the interface. After that, a comparison is performed for each captured image to verify that either the captured meter reading matches with the manually entered reading or not. The mismatched images are given classification codes, such as blur, not readable, no image, and wrong image. Moreover, it is not practically possible to audit the millions of consumers' records before the issuance of bills to the consumers; hence, the images are randomly selected through stratified sampling, which is less than 1% of the total images.

This manual auditing process is time-consuming and requires a lot of human effort as they have to verify millions of records. Also, it can lead to negligence and errors. Furthermore, this system does not allow the observation of real-time statistics for analysis on a weekly or monthly basis. Underbilling and overbilling are also major problems with this system. Hence, there is a need for a robust system that can not only minimize

human effort and cost but also read the meter reading images with more accuracy.

3.2. New System Design. The development of a new system is significantly important because of the major problem of the existing system working at large scale and dealing with millions of consumers' records for billing, troubleshooting, or analyzing purposes. In our new system design, we are replacing the manual auditing process with a novel offline OCR approach. Auditing all the images using the manual approach is devastating in terms of time and cost, in addition to resource consumption. We design a new end-to-end AMR system that incorporates a unified approach to (i) reduce manpower and cost, (ii) improve recognition results (especially in multiple scenarios), and (iii) significantly reduce the number of meter reading images for manual verification by filtering out the images that present illegible/defective meters.

In this study, we develop two web applications: one is the local application, and the one is the main application. The local application is based on an OCR approach that will be used by the service company's employees. This OCR-based application extracts the readings from the captured images and compares them with the manual readings stored by a service company employee (Figure 4). The matched images are stored in a database with a status flag of "matched," whereas unmatched images are saved with an "unmatched" status. For each mismatched image, our system provides a reason why it could not be matched, whether it is blurry or has high illumination. This system also ensures whether captured images are meter images or not. Another feature of the local application is that it pushes all the statistics about meter reading audits to the main application at regular intervals (Figure 5).

The main application has been designed for higher authorities to see the results on a dashboard pushed by the local application on a daily or monthly basis, as shown in Figure 6. This application also provides real-time analysis of the data, whereas the existing system does not incorporate such an attribute for the processed data.

3.3. Novel Offline OCR Approach. This section provides details about novel offline OCR based on machine learning algorithms. This offline OCR is part of a local application that is developed to perform manual auditing process and minimizes human effort and cost. In the first step, we need to localise the box containing the digits in the meter reading image. After that, we classify each digit with the correct label, 0–9. For this purpose, YOLOv5 [20] is employed to accomplish the meter reading detection process. As a result, we obtain a meter reading segment, which is our region of interest (ROI). In this ROI, YOLOv5 detects each classified object. Finally, we

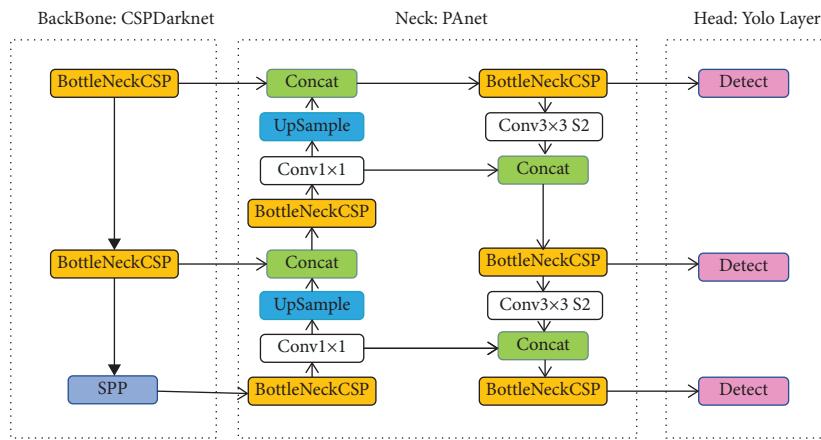


FIGURE 2: YOLOv5's default structure: the data are first input to CSPDarknet for feature extraction and then fed to PANet for feature fusion. Finally, the Yolo layer outputs detection results (class, score, location, and size).

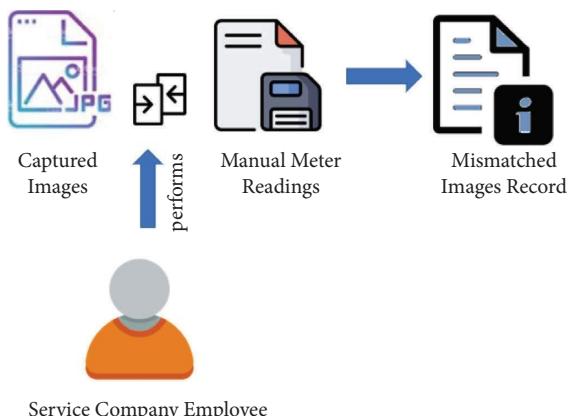


FIGURE 3: Existing system structure.



FIGURE 4: New system design.

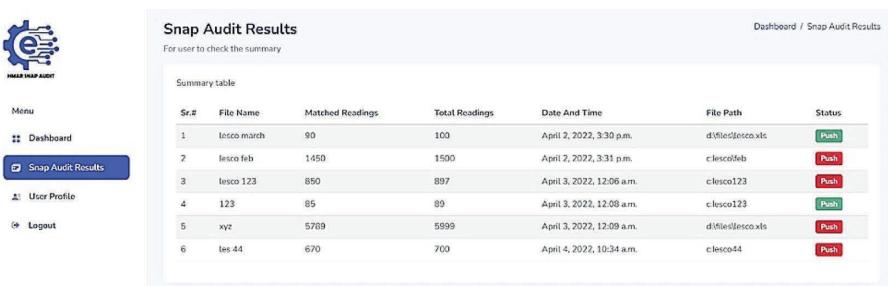


FIGURE 5: Local application for auditing processes.



FIGURE 6: Dashboard design with real-time statistics.

apply a sorting algorithm to read the meter reading in the same order as reading the image. The reason is that existing YOLOv5 returns the detected meter reading numbers in a random manner. It places the number at the first place with shortest distance to the boundary box. To address this issue, we employ a sorting algorithm that returns the result in the same order as the captured meter reading. We also generate a back-end file where each reading is saved against a meter reading image file name. The pseudo-code for this approach is given in Algorithm 1. The flow of the developed approach is shown in Figure 7.

4. Experimental Analysis

4.1. Experimental Setup

4.1.1. Real-World Dataset. In this section, the experiments are conducted on a broad range of 5,000 real-world data images. These meter reading images are taken by a service company employee who goes house to house and building to building on a monthly basis and captures meter snaps, which are later stored on a centralized server.

The main challenge of the dataset is the image quality. It can be affected by various factors, e.g., low-end cameras, complex environmental conditions, and high compression. The challenging environmental conditions include rain, dirt, and low illumination, which results in noisy, blurred, and low-contrast images. Some samples are shown in Figure 8. It can be observed that only the images in Figures 8(b), 8(f), and 8(g) show clear and readable meter readings, whereas others are blurry, are unreadable, or have a high illumination effect.

4.1.2. Parameters. Table 1 shows the hyperparameters of the YOLOv5 used in the experiments.

4.1.3. Experimental Environment. In our experiments, we compile and run YOLOv5 on the Collaboratory. All the experiments are implemented in Python 3.9 with PyTorch using Google cloud disc GPU, running on a personal computer with an Intel Core i5, 3.2 GHz CPU, and 16 GB of RAM.

4.1.4. Performance Metrics. In this section, we present common quantitative metrics to evaluate the performance

of classifiers. In machine learning classification problems, the confusion matrix is a popular performance measurement where the output is two or more classes [37]. We can define it in the form of a table with combinations of predicted and actual values, as shown in Figure 9. Our problem is binary classification, where the meter reading is matched or not.

Another important evaluation metric is accuracy, which is derived from the confusion matrix and is one of the basic measures. It simply measures how often the classifier predicts correctly. Accuracy (1) is the ratio of the number of correct predictions to the total number of predictions.

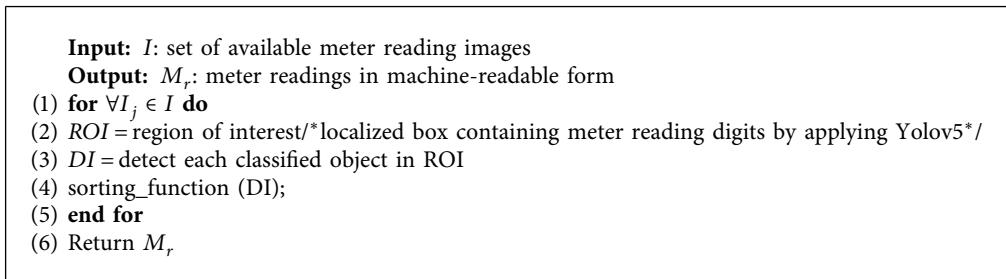
$$\text{Accuracy} = \frac{TP + TN}{FP + TN + TP + FN}. \quad (1)$$

4.2. Results and Discussion. In this section, the performance of the developed approach against state-of-the-art algorithms is presented. The main purpose of this comparison is to validate that our developed offline OCR approach impacts and reduces human time and cost for meter reading verification, i.e., the manual auditing process. To address this, we compare the results of the developed approach with those achieved by existing approaches in Table 2. The results show that our developed approach processes the image in a shorter time with significant accuracy as compared to other algorithms.

From the result of processing the 5,000-image data, YOLOv5 returns with the shortest duration, 0.05 seconds per image, with 94% accuracy. RCNN runs for the longest duration (0.8 seconds), whereas the accuracy rate is 92%. Output duration and accuracy of the nanonets are closer to YOLOv5, but they are not open source. Fast-OCR and PyTesseract have the worst accuracy rates at 10% and 15%, respectively. Figure 10 shows different training and validation results along with different metrics. This figure presents three different types of losses known as box loss, objectness loss, and classification loss. The algorithm's efficiency in locating the centre of an object and how accurately the predicted bounding box covers an object is determined by box loss. Objectness represents the probabilistic measure of whether an object exists in a proposed ROI or not. The high objectivity value means that the object is likely to be in the image window. On the other hand, the algorithm's efficiency in predicting the correct class of the object is measured by classification loss. It can be observed



FIGURE 7: Developed novel offline OCR approach flow.



ALGORITHM 1: Novel offline OCR algorithmic approach.

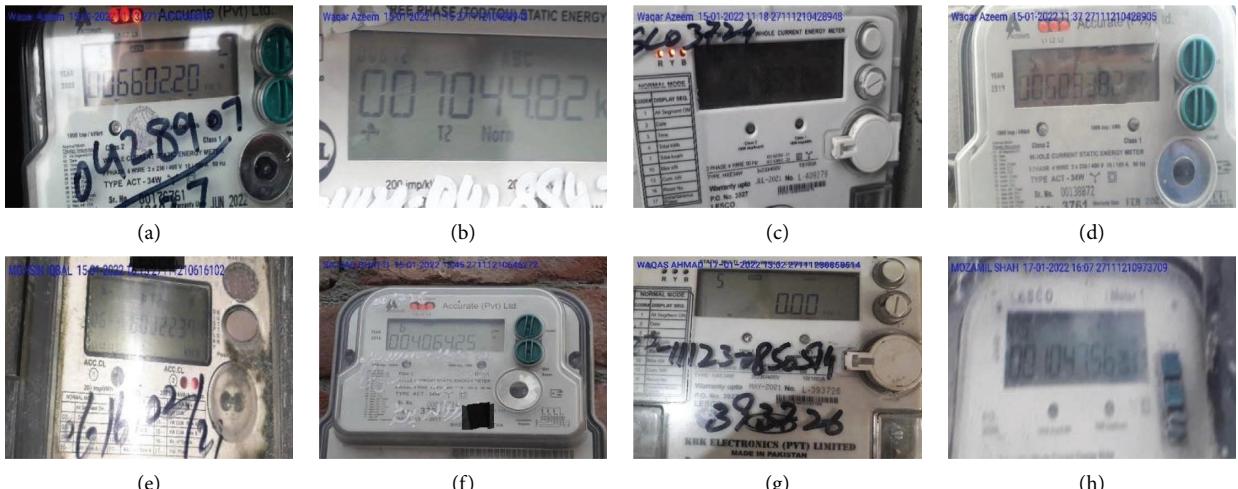


FIGURE 8: Real-world images, taken in a complex environment. (a) Real-world sample 1. (b) Real-world sample 2. (c) Real-world sample 3. (d) Real-world sample 4. (e) Real-world sample 5. (f) Real-world sample 6. (g) Real-world sample 7. (h) Real-world sample 8.

TABLE 1: Hyperparameters of the YOLOv5.

| Parameters | Values |
|-----------------------------|-------------|
| Input size | 224 * 224 |
| Batch size | 64 |
| Learning rate | 0.001 |
| Momentum | 0.9 |
| Train/validation/test split | 70%/20%/10% |

from the results that, for precision, recall, and mean average precision metrics, the model improved swiftly after epoch 50. There is also a rapid decline until around epoch 50 in the box, and objectness losses of the validation data are observed.

To further validate the effectiveness of our approach, the results of state-of-the-art algorithms and our approach are shown in Figure 11. The results demonstrate that nanonets only detect the meter reading region without giving any further details. For Fast-OCR, only partial reading is detected, whereas PyTesseract is not able to detect and read

| | | Predicted Class | |
|------------|---------------------|--------------------|---------------------|
| | | True Positive (TP) | False Negative (FN) |
| True Class | False Positive (FP) | True Negative (TN) | |
| | | | |

FIGURE 9: Confusion matrix.

TABLE 2: Performance comparison of developed vs. existing algorithms.

| Algorithm | Time | Cost | Working mode | Accuracy (%) |
|-------------|------|-------------|--------------|--------------|
| YOLOv5 | 0.05 | Open source | Offline | 94 |
| RCNN | 0.8 | Open source | Offline | 92 |
| Nanonet | 0.73 | Paid | Online | 95 |
| Fast-OCR | 0.4 | Open source | Offline | 10 |
| PyTesseract | 0.1 | Open source | Offline | 15 |

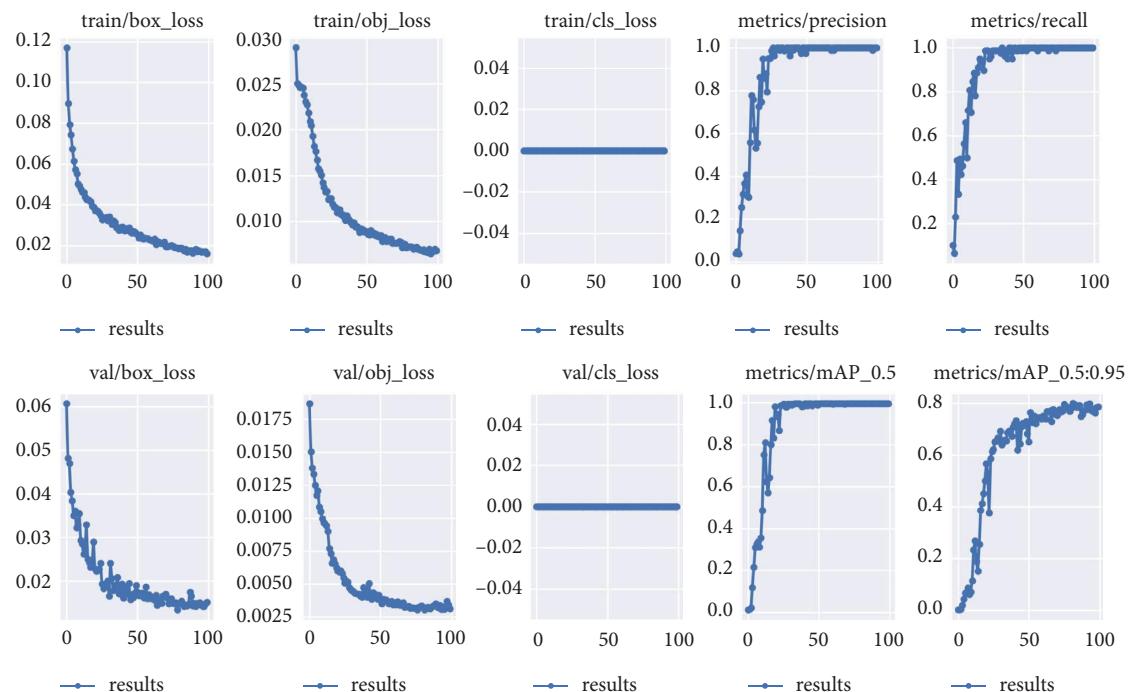


FIGURE 10: Box loss, objectness loss, classification loss, precision, recall, and mean average precision (mAP) plots for the training and validation set.

meter reading correctly. Another interesting observation is that RCNN detects the region boundary but with zero recognition rate. In this regard, YOLOv5 shows distinct superiority over its peers and detects the meter reading with a higher recognition rate. In Figure 12, it can also be observed that our developed offline OCR approach recognises each digit in a meter reading image correctly.

5. Threats to Validity

This study aims to reduce the human effort and cost of the manual meter reading auditing process in developing countries. However, there are some limitations to this work. First, this study is only valid in those places where the meter reading process is manual and requires a lot of human effort and time. In this study, we have evaluated a dataset with 5,000 real-world images. Resiliency is another problem in a



FIGURE 11: Comparison results of state-of-the-art algorithms on real-world meter reading image data. (a) Nanonets result 1. (b) Nanonets result 2. (c) Fast-OCR result. (d) PyTesseract result 1. (e) PyTesseract result 2. (f) RCNN result 1. (g) RCNN result 2. (h) YOLOv5 result 1. (i) YOLOv5 result 2. (j) YOLOv5 result 3. (k) YOLOv5 result 4. (l) YOLOv5 result 5. (m) YOLOv5 result 6. (n) YOLOv5 result 7. (o) YOLOv5 result 8.

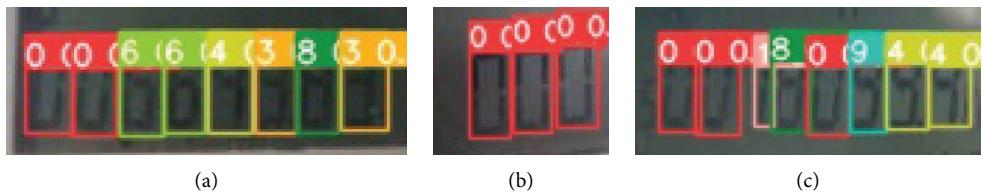


FIGURE 12: Developed OCR results. (a) Meter reading: 006643830. (b) Meter reading: 000. (c) Meter reading: 001809440.

complex environment. It means that the solution generated by our approach is 100% identical to that read by a human computer operator.

6. Conclusion and Future Developments

Due to the high cost involved with procurement, installation, and operationalization of smart meters, developing countries are still using analogue or digital meters. The human effort required to ensure meter reading accuracy for each billing cycle through manual meter reading and subsequent image auditing is huge, and this process cannot be performed with 100% accuracy due to limited resources. In this study, we present a new system design that aims to reduce the human effort and cost without any structural change in the process or additional hardware inclusion, as well as the capability to carry out 100% offline auditing. We have proposed a novel offline OCR for effective meter reading with high accuracy. The developed approach is evaluated with four state-of-the-art algorithms using a 5,000-image dataset, some of them taken from complex environments. The comparative analysis demonstrates that the newly developed offline OCR approach is effective in terms of processing time, cost, and accuracy and offers extended 46.3% accuracy in comparison with state-of-the-art OCR algorithms. Based on these facts, the new system design and developed approach will be very helpful for service companies in developing countries, in terms of human effort and cost, with more accurate results.

In future work, the developed approach will be evaluated with more data images, taken in both normal and complex environments.

Data Availability

Data are available on request to the corresponding author.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Hafiz Muhammad Faisal, Muhammad Kashif Shahzad, and Shahid Islam contributed equally to the study.

References

- [1] R. Laroca, V. Barroso, M. A. Diniz, G. R. Gonçalves, W. R. Schwartz, and D. Menotti, "Convolutional neural networks for automatic meter reading," *Journal of Electronic Imaging*, vol. 28, no. 1, Article ID 013023, 2019.
- [2] R. C. da Silva Marques, A. C. Serra, J. V. F. França et al., "Image-based electric consumption recognition via multi-task learning," in *Proceedings of the 2019 8th Brazilian Conference on Intelligent Systems (BRACIS) IEEE*, pp. 419–424, Salvador, Brazil, October 2019.
- [3] L. Meng and J. Cheng, "Research on the visual recognition method of pointer water meter reading," in *In Proceedings of 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, pp. 909–914, Chongqing, China, March 2021.
- [4] M. Vanetti, I. Gallo, and A. Nodari, "Gas meter reading from real world images using a multi-net system," *Pattern Recognition Letters*, vol. 34, no. 5, pp. 519–526, 2013.
- [5] I. Gallo, A. Zamberletti, and L. Noce, "Robust angle invariant GAS meter reading," in *Proceedings of the 2015 International Conference on Digital Image Computing: Techniques and Applications (DICTA) IEEE*, pp. 1–7, Adelaide, SA, Australia, November 2015.
- [6] M. Waqar, M. A. Waris, E. Rashid, N. Nida, S. Nawaz, and M. H. Yousaf, "Meter digit recognition via Faster R-CNN," in *Proceedings of the 2019 International Conference on Robotics and Automation in Industry (ICRAI) IEEE*, pp. 1–5, Montreal, Canada, May 2019.
- [7] M. Cerman, G. Shalunts, and D. Albertini, "A mobile recognition system for analog energy meter scanning," in *Proceedings of the International Symposium on Visual Computing*, pp. 247–256, Las Vegas, NV, USA, December 2016.
- [8] Y. Kabalci, "A survey on smart metering and smart grid communication," *Renewable and Sustainable Energy Reviews*, vol. 57, pp. 302–318, 2016.
- [9] R. Zhang, Y. Yang, and W. Wang, "Research on document digitization processing technology," *MATEC Web of Conferences*, vol. 309, Article ID 02014, 2020.
- [10] P. K. Singh, S. V. Singh, M. K. Singh et al., *International Journal of Immunogenetics*, vol. 39, no. 4, pp. 314–320, 2012.
- [11] R. Mithe, S. Indalkar, and N. Divekar, "Optical character recognition," *International Journal of Recent Technology and Engineering*, vol. 2, no. 1, pp. 72–75, 2013.
- [12] A. Lat and C. Jawahar, "Enhancing OCR accuracy with super resolution," in *Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR) IEEE*, pp. 3162–3167, Beijing, China, June 2018.
- [13] X. Peng and C. Wang, "Building super-resolution image generator for OCR accuracy improvement," in *Proceedings of the International Workshop on Document Analysis Systems*, pp. 145–160, Boston, Massachusetts, June 2020.
- [14] C. Li, Y. Su, R. Yuan, D. Chu, and J. Zhu, "Light-weight spliced convolution network-based automatic water meter reading in smart city," *IEEE Access*, vol. 7, Article ID 174359, 2019.

- [15] K. Kanagarathinam and K. Sekar, "Text detection and recognition in raw image dataset of seven segment digital energy meter display," *Energy Reports*, vol. 5, pp. 842–852, 2019.
- [16] N. Nigar, M. K. Shahzad, S. Islam, S. Kumar, and A. Jaleel, "Modeling human resource experience evolution for multi-objective project scheduling in large scale software projects," *IEEE Access*, vol. 10, pp. 44677–44690, 2022.
- [17] L. Zuo, P. He, C. Zhang, and Z. Zhang, "A robust approach to reading recognition of pointer meters based on improved mask-RCNN," *Neurocomputing*, vol. 388, pp. 90–101, 2020.
- [18] K. Košćević and M. Subašić, "Automatic visual reading of meters using deep learning," in *Croatian Computer Vision Workshop*, pp. 1–6, 2018.
- [19] C. M. Tsai, T. D. Shou, S. C. Chen, and J. W. Hsieh, "Use SSD to detect the digital region in electricity meter," in *Proceeding of the 2019 International Conference on Machine Learning and Cybernetics (ICMLC)*. IEEE, pp. 1–7, Kobe, Japan, March 2019.
- [20] J. Glenn, <https://github.com/ultralytics/yolov5>, 2020.
- [21] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, Canada, June 2014.
- [22] S. Online, <https://nanonets.com/ocr-api>, 2020.
- [23] R. Laroca, A. B. Araujo, L. A. Zanlorensi, E. C. De Almeida, and D. Menotti, "Towards image-based automatic meter reading in unconstrained scenarios: a robust and efficient approach," *IEEE Access*, vol. 9, Article ID 67569, 2021.
- [24] S. Online, <https://pypi.org/project/pytesseract/>, 2022.
- [25] S. Marinai and H. Fujisawa, *Machine learning in document analysis and recognition*, Vol. 90, Springer, , Heidelberg, Germany, 2007.
- [26] L. Gómez, M. Rusiñol, and D. Karatzas, "Cutting sayre's knot: reading scene text without segmentation. Application to utility meters," in *Proceedings of the 2018 13th IAPR International Workshop on Document Analysis Systems (DAS)*, pp. 97–102, Vienna, Austria, June 2018.
- [27] F. Yang, L. Jin, S. Lai, X. Gao, and Z. Li, "Fully convolutional sequence recognition network for water meter number reading," *IEEE Access*, vol. 7, Article ID 11679, 2019.
- [28] G. Salomon, R. Laroca, and D. Menotti, "Deep learning for image-based automatic dial meter reading: dataset and baselines," in *Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, pp. 1–8, Glasgow, UK, Januray 2020.
- [29] G. Salomon, R. Laroca, and D. Menotti, "Deep learning for image-based automatic dial meter reading: dataset and baselines," in *Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, San Francisco, Canada, March 2020.
- [30] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," 2015, <https://arxiv.org/abs/1511.08458>.
- [31] J. R. R. Uijlings, K. E. A. van de Sande, T. Gevers, and A. W. M. Smeulders, "Selective search for object recognition," *International Journal of Computer Vision*, vol. 104, no. 2, pp. 154–171, 2013.
- [32] G. Kumar and P. K. Bhatia, "A detailed review of feature extraction in image processing systems," in *Proceedings of the 2014 Fourth International Conference on Advanced Computing Communication Technologies*, pp. 5–12, Washington, DC, USA, July 2014.
- [33] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 779–788, Las Vegas, NV, USA, 2016.
- [34] A. Benjumea, I. Teeti, F. Cuzzolin, and A. Bradley, "YOLO-Z: Improving Small Object Detection in YOLOv5 for Autonomous Vehicles," 2021, <https://arxiv.org/abs/2112.11798>.
- [35] C. Y. Wang, H. Y. M. Liao, Y. H. Wu, P. Y. Chen, J. W. Hsieh, and I. H. Yeh, "CSPNet: a new backbone that can enhance learning capability of CNN," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 390–391, Long Beach, CA, USA, June 2020.
- [36] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8759–8768, Silver Spring, MD, USA, December 2018.
- [37] S. Visa, B. Ramsay, A. L. Ralescu, and E. Van Der Knaap, "Confusion matrix-based feature selection," *MAICS*, vol. 710, no. 1, pp. 120–127, 2011.