2021 International Conference on
Artificial Intelligence (ICAI)
Islamabad, Pakistan, April 05-07, 2021

Brick Kiln Detection and Localization using Deep Learning Techniques

Rosheen Tahir
Department of Computer Science
Forman Christian College
Lahore, Pakistan
rosheentahir@gmail.com

Muhammad Shaaf Imran
Department of Computer Science
Forman Christian College
Lahore, Pakistan
shaaf19@live.com

Sidra Minhas
Department of Computer Science
Forman Christian College
Lahore, Pakistan
sidraminhas@fccollege.edu.pk

Nosheen Sabahat
Department of Computer Science
Forman Christian College
Lahore, Pakistan
nosheensabahat@fccollege.edu.pk

Sardar Haider Waseem Ilyas
Suleman Dawood School of Business
Lahore University of Management Sciences
Lahore, Pakistan
haider@waseemilyas.com

Haider Raza Gadi
Department of Computer Science
Forman Christian College
Lahore, Pakistan
haider.raxa23@gmail.com

2021 International Conference on Artificial Intelligence (ICAI) | 978-1-6654-3293-1/20/$31.00 ©2021 IEEE | DOI: 10.1109/ICAI52203.2021.9445267

Abstract*—Brick kiln areas being home to practices of forced labor and environmental pollution are difficult to detect. This paper explores the automatic detection and localization of brick kilns in satellite images by the implementation of Deep Learning Techniques to aid the government to take action against forced labour and environmental pollution. An artificially intelligent model was trained to automatically detect and create a bounding box around the kilns in satellite images from the South Asian brick belt. The experimentation started with the use of basic convolutional neural networks to identify kilns in an image and then progressed to the more advanced You Only Look Once (YOLOv3) algorithm to draw bounding boxes across the kilns. Our research enabled us to successfully build a convolutional neural network for the identification of kilns which returned an accuracy of 97.27%. For the localization of kilns using the bounding box approach, our research enabled us to implement the YOLOv3 algorithm which returned an average loss of 0.13 and an average IOU of 0.75. The work done in this project has been promising thus far and can be further enhanced to build a Graphical User Interface to help governments to take action against the social issues of bonded labor and environmental issues that are prevalent within brick kilns.

***Keywords-** artificial intelligence, deep learning, convolutional neural network, You Only Look Once, brick kilns, satellite images, forced labor, environmental pollution.*

I. Introduction

The manufacturing of bricks is a multi-stage process which includes preparation of clay, molding, drying and burning of bricks. According to a World Bank report [1], South Asia caters to about a quarter of brick production in the world. Bangladesh is estimated to produce about 27 billion bricks per year and employs nearly 1 million people. In Pakistan, brick kilns are the main source for the production of bricks and according to the Labour and Human Resource Department of Punjab [2], there are approximately 8000 kilns in Punjab and 18000 in the whole country. Bricks are primarily used for the construction of basic structure of buildings. They are also useful in making concrete as broken pieces of bricks are an aggregate of concrete. Similarly, powdered bricks are used in the manufacture of lime plaster and lime concrete. Thus, bricks are an integral part of the construction industry in Pakistan.

The brick belt in South Asia which runs through Afghanistan, Pakistan, India, Bangladesh and Nepal consists of many traditional brick kilns which are used in the production of bricks. The manufacturing process is highly labor intensive and kiln owners often use forced labor on the manufacturing site. The term forced labor is coined to include child labor, enslaved workers and all forms of bonded labor.

The Global Slavery Index of 2018 shows how 24.9 million people are a part of forced labor worldwide [3]. Out of all the people trapped in forced labor globally, 12.7 million people are estimated to reside within the brick belt in South Asia [4]. The UN's Sustainable Development Goal (SDGs) 8 aims to eradicate modern slavery including forced labor by the year 2030 [5]. However, they are limited by the lack of access to reliable data to locate areas of slavery. Brick Kiln areas are one of the areas that have a large number of enslaved workers and are not outlined on any of the most widely used mapping services, therefore, the need arises to find an automated solution to locate these areas.

Another aspect of brick kilns involves the analysis of its impact on the environment and human health. According to a study conducted in India [6], the sources of air pollution and the impact of air pollution on human health and crops was examined. Compounds of Sulfur and compounds of Carbon were identified as the major pollutants in the atmosphere. Brick kilns were then identified as one of the major source of these pollutants. Research on the harmful effects of these pollutants show that emissions from brick kilns disrupt lung function, causes narrowing of airways and give rise to breathing problems. WHO has also stated that the pollutants in the atmosphere are the cause of about 21,000 premature deaths per year in 25 countries in the WHO European Region [7]. Moreover, the air pollutants released from kilns were also seen to deteriorate fertility of agricultural land. In another study, the concentrations of ozone and nitrogen dioxide emitted from brick kilns were examined [8]. The experiments concluded that brick kilns are a point source of the precursor gases of ozone with high levels being recorded during the periods when kilns are active. In underdeveloped countries such as Pakistan and Bangladesh, smoke emissions from brick kiln industries are a major cause of the development of smog. The levels of smog in the air has increased over the years which in turn give rise to a number of health problems such as asthma and breathing disorders. Looking at the issue of smog from the perspective of hazardous air quality in Pakistan, the government of Pakistan has ordered a shutdown of kilns for a month to combat the issue of smog [9]. However, the kilns are scattered across the country and are large in number about 19,000 which means that even if the government passes a law to control the operation of kilns, there is no definite way of locating and monitoring the kilns. Brick kilns are scattered across the world and one way to identify them is by using satellite imagery from Google Earth.

Keeping the above-mentioned issues in view, there is a need for an automated solution to detect and locate brick kilns. This study proposes an automated solution for the detection and localization of kilns. The research carried out was divided into two parts: (i) Detection of Kilns and (ii) Localization of kilns. We applied our proposed algorithm on self-developed satellite images from the South Asian brick belt with a focus on the region of Bangladesh and a few images from Pakistan used in testing. For detection, Python was used to create a Convolutional Neural Network (CNN) to identify whether a picture contains or does not contain a brick kiln. Similarly, for localization, our research proposed the use of image augmentation and annotation techniques combined with the use of You Only Look Once (YOLOv3) algorithm to create bounding boxes around kilns. The findings of our research can be further enhanced and used by the government and Non-Governmental organizations to locate brick kiln areas to fight against forced labour and environmental pollution.

II. Related Work

The availability of high-resolution remote sensing images coupled with advancements in deep learning techniques has paved the way for researchers to carry out object detection using various algorithms. The topic of brick kilns has been discussed across multiple works of literature including the use of bonded labour on kilns and environmental concerns due to smoke emissions from the kilns. Research on using satellite images to address these social and environmental issues has been carried out using different datasets and different algorithms.

The idea of using satellite images was introduced by a crowdsourcing project [10] in which volunteers manually detected kilns in satellite images. The research led to the creation of a dataset of kilns to be used later in the study of machine learning algorithms. The study focused on the first phase of kiln identification which was the creation of a dataset. Similarly, in another study [3] expert visual representation was used to identify kilns in satellite images. The method incorporated the use of standard deviation to estimate kilns in a region along with the use of volunteers via a crowd-sourcing platform to identify brick kilns in the area of Rajasthan. Volunteers analyzed image extracts presented randomly from the 396 image extracts that covered the region. Each image extract was viewed and annotated by approximately 15 volunteers to aid quality assessment. The image data for each of these selected grid cells were visually interpreted to get a count of the number of kilns present in the area. The average accuracy attained by these studies was approximately 96% and the studies adopted a highly labour-intensive approach for the said purpose.

Machine learning methods were then introduced to automate labour-intensive methods. In a study carried out at the Carnegie Melon University [11], the Support Vector Machine (SVM) was combined with the nearest neighbour approach to perform object detection and classification on the PASCAL VOC-2007 dataset. The experiment results yielded an average mAP of 27%.

Moving forward, Neural Networks were introduced to create different object detectors to overcome the shortcomings of the previously used support vector model. The addition of backpropagation, multiple hidden layers, batch normalization and activation features made neural networks a more preferred model for object detection. The use of Pretrained weights in CNN introduced the concept of transfer learning for object prediction and detection. In a study conducted by researchers at the Stanford University [12] transfer learning and prediction of night-time light intensity was used to extract indicators of poverty from images. The model used in the study performed feature extraction from high-resolution images. These features include man-made structures, roads, buildings, farmlands which are then further analysed to locate areas of poverty. These indicators of poverty were used to identify regions surrounding brick kilns where bonded labour and poverty levels are known to be high. The transfer learning approach combined with a CNN resulted in a model for automatic detection which returned an accuracy of 71.6%.

Similarly, in another research conducted in India [5] brick kilns were identified using the Faster R-CNN model for detection. The initial experiments returned an accuracy of 100% but at the cost of an overestimation of kilns coupled with high commission errors. Therefore, a second classifier was added to the model to reduce commission errors. The final accuracy obtained after using the Faster R-CNN combined with a CNN as a second classifier obtained was 94.94%.

Our approach of using a CNN for detection of brick kilns contributed to previous studies in the domain by providing an architecture that was implemented using Google Colaboratory and returned an accuracy of 97.27%. The use of Jupyter Notebooks on Google Colaboratory makes our approach widely accessible to researchers who are challenged with low computing power resources.

To take this research one step further, the second part of our study focused on the automatic localization of kilns in satellite images. The concept of localization was initially implemented in a study that carried out automatic target detection on a military database to draw boxes around aeroplanes in images [13]. The boxes were drawn using the Edge Boxes algorithm for object detection and CNN for classification in satellite images. After testing, the experiments returned an average precision of 77.9% and an average recall of 91.3%.

Similarly, another multi-scale CNN for geospatial object detection was built to create bounding boxes in high-resolution satellite images [14]. The NWPU VHR-10 dataset with ten different classes was used for object detection. A comparison with seven other object detection algorithms yielded the highest mean average precision of 89.6% for their proposed architecture. The comparison of the mean average precision (mAP) of Faster R-CNN and SSD gives an insight into the functionality of other object detection algorithms. The multi-scale CNN returned a good mAP as the algorithm was able to detect both large and small objects. However, some bounding boxes were seen to be repeated multiple times which lead to the further need for improvement in accuracy for localization.

In the study of single-shot multi-box detectors, the concept of the You Only Look Once (YOLO) algorithm was introduced in 2016 [15]. Object detection was framed as a regression problem and an algorithm was created to process the whole image at one instance. The performance of YOLO was compared with R-CNN using the Pascal VOC 2007 dataset. For further analysis, two artwork datasets Picasso and People-Art were used to show that YOLO generalizes better on new domains as compared to other object detection algorithms. YOLO returned the highest average precision of 53% in all three datasets. Along with room for improvement in terms of precision, YOLO showed problems in the detection of different sizes of input. If the training was done on small images of a specific object, the architecture showed problems detecting the same images on a larger size.

The drawbacks of YOLO led to another study [16] which built on to the YOLO architecture to create a better object detection algorithm. The base feature extractor of YOLO was compared with other object detectors. For most object detectors the base feature extractor used was VGG-16 but YOLO used a custom architecture based on GoogleNet architecture. The comparison with VGG-16 showed that the custom GoogleNet architecture yielded a lower accuracy. To improve the accuracy and performance of YOLO, a new architecture known as DarkNet-19 was introduced as the base for YOLOv2 which contained 19 layers of convolution and 5 layers of max pooling. Similarly, in another paper, Dixit [17] analyses the upgrades to YOLO which included the addition of batch normalization, the increase in input size for high-resolution images, the addition of anchor boxes and fine-grained features were analyzed all of which boosted the mean average precision. The study compared the performance of YOLO and YOLOv2 on the PascalVOC-2007 dataset which returned a mean average precision value of 63.4 % for YOLO and 78.6% for YOLOv2.

Moving forward, a stronger architecture known as YOLOv3 was introduced in 2018 [18]. The base feature extractor of this model improved from DarkNet-19 to a 53 layer network called DarkNet-53 which helped to improve the accuracy of the model. This model used residual skip connections, upsampling, and made detections at three different scales. 1x1 detection kernels were applied to feature maps of three different sizes at three different places in the network. This algorithm was able to predict a greater number of bounding boxes than previous versions which also increased the precision of the architecture. The performance of multiple variants of YOLO [18] along with other single-shot detectors [19] were compared on the COCO dataset which returned YOLOv3 as the model with the highest mean average precision of 57.9%.

After evaluating the different literature on object detection algorithms and looking at previous studies on brick kiln detection, our approach contributes by implementing the YOLOv3 with satellite images to locate kilns in satellite images. Access to the code on Google Colaboratory makes it widely accessible for implementation by researchers that do not have access to a strong GPU. Further enhancements to our research could be used to develop a system to locate brick kilns to aid action against forced labour and environmental pollution.

III. Research Methodology

Our proposed study was divided into two major parts: Detection and Localization. Fig. 1 shows the flow diagram of the research methodology that was implemented for the detection and localization of kilns in satellite images.

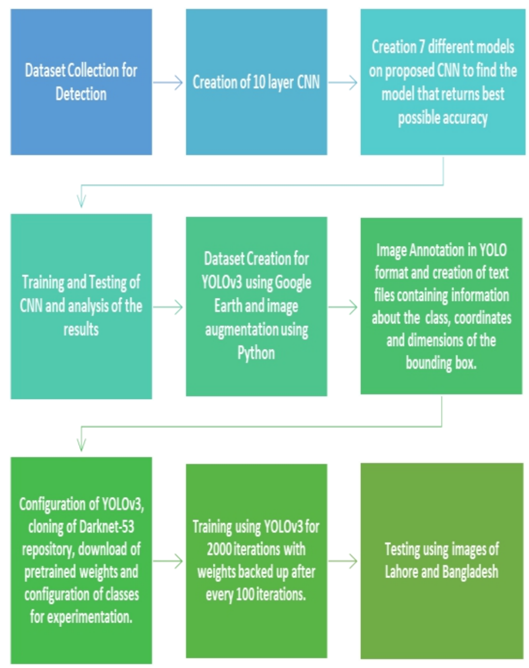

Fig. 1: Flow Diagram of Proposed Study

A. Dataset Description

We used one dataset for detection and a different dataset for localization. We used two different datasets because the one used for detection had multiple images that were magnified and drawing a bounding box on those images meant drawing a bounding box on the whole image. Therefore, to increase the variety of images and cater images with multiple kilns we created a new dataset for localization. Dataset collection for detection was done with the help of a local institution in Pakistan. The dataset consisted of 4000 images with and without kilns each having dimensions of 256x256. For the implementation of the bounding box approach, we developed our dataset (Fig. 2) by gathering satellite images from Google

Earth. A total of 200 images of kilns in Bangladesh were collected from different angles and heights on Google Earth. The dimensions of the pictures were 1310x460 but when entered into YOLOv3 all images were resized into 416x416 dimensions which was the default size for this algorithm. To increase the dataset, image augmentation techniques via Python were used to create a final dataset of 600 images.

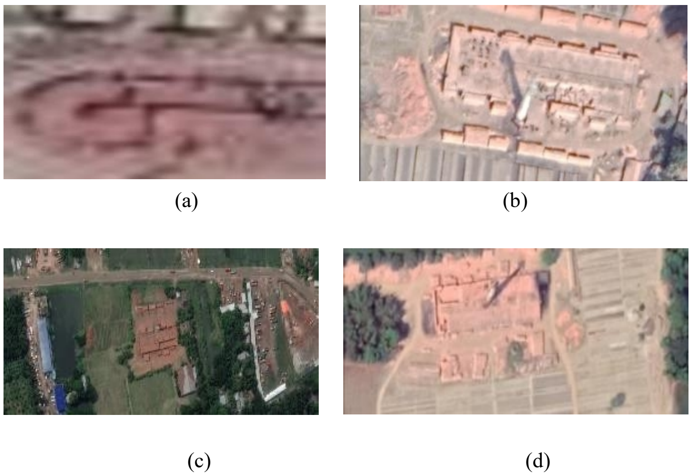

Fig. 2:Sample image for detection (a&b), sample image for localization (c&d).

B. Convolutional Neural Network (CNN) for Detection of Brick Kilns

A multi-layer CNN consisting of 10 layers was built using the convolution layer, pooling and dense layers to which weights were applied to predict labels for the images in the dataset. Images were fed sequentially to the algorithm. The dataset was trained six times with modification of parameters in each experiment. A total of seven experiments were conducted to attain the highest accuracy possible. The evaluation of the model for kiln detection was done using the Accuracy formula given below:

$$Accuracy = \frac{(Predicted\ True)}{(Actual\ True)} \tag{3.2}$$

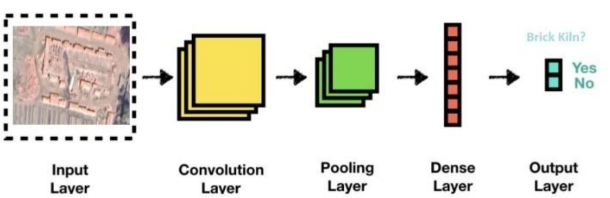

Fig. 3: Simplified Architecture of CNN used for detection

1) Experiment 1:

In the first experiment, a single-layered model with one layer of convolution was used along with one layer of pooling and further single layers of flattening and densing of the neural network. 10 filters were set along with a kernel size of 3x3 with stride of 2. The network was densed down to 2 (resulting in a yes or no format) and the cycles were set to 5 cycles(epochs).

2) Experiment 2

In the second experiment, a second layer of convolution was added to the model. The first layer was modified to learn 64 filters, and the second was made to learn 128 filters. The other parameters remained the same.

3) Experiment 3

In the third experiment, 2 layers of densing were added to the model. Densing was done stepwise, first a layer of 1024 was added which was reduced to 256 and then further reduced to 2. The other parameters remained the same as the previous experiment.

4) Experiment 4

During the fourth experiment, Batch normalization (Backpropagation) was added to the model.

5) Experiment 5&6

In the fifth and sixth experiment, the number of epochs was increased from 5 to 20 to 50, which were used to stabilize the results.

6) Experiment 7

The seventh and last experiment had a change in the learning rate of the model. Initially, the learning rate was set at 1 but it was modified and changed to 0.00001. The learning rate is a hyperparameter that controls how much to change the model in response to the estimated error each time the model weights are updated. Large learning rates resulted in unstable and varying training results as seen above therefore lowering the learning rate drastically helped to stabilize and improve the results of our model.

C. Testing the CNN model for Detection

The model from experiment 7 was saved and used to test whether it could identify a brick kiln in a satellite image or not. The basic output from the model was a number near 0 or 1. A 0 depicted the existence of a brick kiln while a 1 depicted the absence of a brick kiln. For testing purposes, an image with a brick kiln was used and the model returned an output of 0.1 which was near zero and stated it belonged to the class of images that contained brick kilns. After creating different models and conducting seven different experimental runs the final results proved to be accurate and successful in predicting whether or not an image contained a brick kiln.

D. Localization with YOLOv3

For YOLOv3, we used our self-developed dataset of images from Bangladesh. The images were annotated in the YOLO format using the labelimg tool. In this format, a text file was created for each image which contained information about the object class, object coordinates, height and width of the bounding box. The images were then mirrored on the x-axis and the y-axis using python for augmentation.

Setting up the YOLOv3 algorithm was a multistage process. The first step was to enable the GPU and begin cloning the DarkNet-53 repository from GitHub. Once cloned, the GPU, OpenCV and CUDA were enabled from the Darknet files which were then integrated into our Colaboratory virtual GPU. The dataset folder was zipped and placed onto Google Drive and every time the code was executed the dataset was unzipped and copied into our data folder in the YOLOv3 algorithm. Changes were made to three files to cater to our custom dataset in the algorithm. The configuration file contained every detail of YOLOv3 including the number of layers and the associated architecture for the implementation of the algorithm. The number of classes was set to 1 which showed the presence of a kiln in an image. The second file for modification contained the name of the classes used for training. A third text file was used in which the number of classes for training, the directories of our files and the backup location was stored. All the files were uploaded on Google Drive. The algorithm was set to run for 2000 iterations and Pretrained weights were downloaded from the YOLO website. The algorithm was set to save the weights after every 100 iterations to prevent data loss in the event of a timeout. Another text file was created to store a list of all the names of the image files intended to be used in training. A python script was used to read the names of the image files and write them on to a text file instantly to be in accordance with the standards of the configuration of YOLO.

Once the algorithm was set up, the images were divided into an 80/20 train test split. The hardware of the machine used consisted of a 2.4GHz Quad-Core i5 processor with 6 gigabytes of RAM. Since all the work was done on Google Colab and its virtual machine, a high spec computer was not needed for our research. Training went up to 2000 iterations in about 6 to 7 hours with the weights backed up on Google Drive after every 100 iterations. Once the training was completed, the algorithm was set into test mode and a combination of images from Lahore and some from Bangladesh were used for testing purposes. The images for testing were uploaded on Google Drive and ran through the YOLOv3 algorithm to create bounding boxes around kilns.

For the evaluation of YOLOv3, the inbuilt loss function along with precision, recall, F1 score and average IoU values were used.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{3.3.1}$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{3.3.2}$$

$$F1\ score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \tag{3.3.3}$$

IV. Results

This section shows the results of our proposed approach.

1) Experiment 1

The first experiment conducted on our model returned a validation accuracy of 65.91% which is displayed in Fig. 4.

2) Experiment 2

The modification of the layers for this experiment in which the first layer was made to learn 64 filters and the second was made to learn 128 filters reduced the accuracy and brought it down to 34.09% .

3) Experiment 3

In the third experiment, the addition of 2 more layers of densing helped to bring back up the accuracy of our CNN to 65.91% .

4) Experiment 4

Batch normalization (Backpropagation) was added to the model in this experiment. In backpropagation, weights are fine-tuned based on the loss attained in the previous epoch. This helped to lower the error rate and made the model more reliable. Addition of batch normalization improved the accuracy to 86.82% (Fig. 5) in the last epoch. However, the results varied in all the epochs.

5) Experiment 5&6

The next two experiments increased the epochs from 5 to 20 and then 50. The accuracy was seen to improve to 87.2%.

6) Experiment 7

The change in learning rate from 1 to 0.00001 helped to stabilize the results and returned a final accuracy of 97.2% (Fig. 6).

7) Results of YOLOv3

Our model ran for 2000 iterations and took about 7 hours to train a set of 600 images. The images in the dataset were divided into an 80/20 train test split. During testing, the execution of our model returned the images within 20 milliseconds. In addition, the probability of a kiln being present in the predicted bounding box was given and the image used for testing appeared with a bounding box around the predicted brick kiln. If there was no brick kiln in an image, no probability was shown and the original image appeared as it is. The mean IoU, average loss, precision, recall and F1 score were calculated for our model. The results are summarized in Table I and the sample images from Lahore, Pakistan before and after running through our proposed algorithm are given in Fig. 7 to Fig. 9.

TABLE I: RESULTS OF YOLOV3

Mean IoU	Average loss	Precision	Recall	F1 Score
0.75	0.13	0.950	0.904	0.926

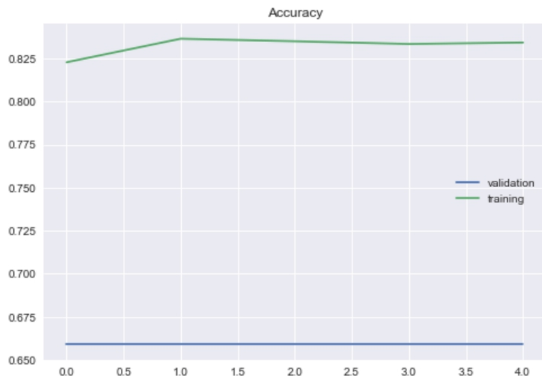

Fig. 4: Results of Experiment 1. The x-axis represents the number of epochs and the y-axis represents the accuracy.

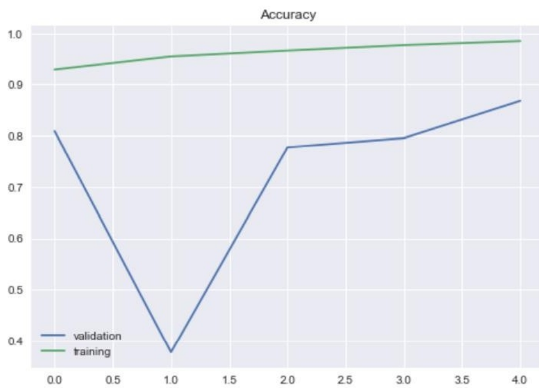

Fig. 5: Results of Experiment 4 after the addition of Batch Normalization. The x-axis represents the number of epochs and the y-axis represents the accuracy.

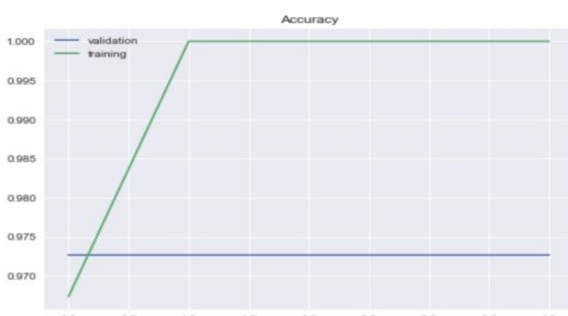

Fig. 6: Results of Experiment 7 after the reduction of Learning Rate with epochs on X axis and accuracy on Y axis.

Fig. 7: Test image of Lahore for YOLOv3

Fig. 8: Test image of Lahore for YOLOv3

Fig. 9: Test image of Lahore for YOLOv3

V. Discussion

Our research revolved around the study of brick kilns in satellite images. Brick kilns are not outlined on any maps (online or offline). The areas surrounding brick kilns are home to a large number of forced labourers and are a source of environmental pollution. To aid action against forced labour and implement laws for environmental protection, our project automated a solution for the automatic detection and localization of kilns using Deep Learning and Image Processing techniques.

For the first part of our research, we focused on the automatic detection of kilns. This involved the automatic detection and differentiation between pictures with and without kilns. The implementation of convolutional neural networks (CNN) was used for the said purpose which returned a final accuracy of 97.2%. The analysis of our approach and comparison with the techniques used in previous studies showed the application of CNN for object detection. When applied to kilns, the studies showed the use of manual methods for the correct identification of kilns. The studies that did apply neural networks used external GPUs which might not be accessible to a larger number of people. Our model showed the implementation of a CNN using a virtual GPU on Google Colab that uses relatively lesser computing power and is accessible to all for future improvements and implementation. In terms of accuracy, our model attained a slightly higher accuracy of 97% as compared to other similar studies in this domain.

After the detection of kilns, we took our research one step further and focused on the localization of kilns. By localization, we meant the ability to identify where in a satellite image is a kiln located. The bounding box approach was used for the said purpose. The images were returned with bounding boxes drawn around the kilns after the implementation of the You Only Look Once (YOLO) algorithm. As compared to R-CNN that analyses images region-wise, this algorithm analyses the whole image at a single instance and creates bounding boxes around the objects. We used the third variant of YOLO as it was the latest one available during our time of research. With the YOLOv3 algorithm, our proposed study was able to compute 2000 iterations in about 6 hours and return an average IoU of 0.75. The concept of bounding boxes has been used in a limited number of studies for brick kilns. Object localization studies done before have used different datasets and used different algorithms. Our research, therefore, contributes to the field by implementing a new algorithm using a virtual GPU that is accessible to all.

The testing of YOLOv3 done on a few images of Lahore shows an avenue for future research and development. Pakistan being home to thousands of brick kilns and the recent increase in hazardous air quality leads to the need for a smart solution to aid the government to take action. Thus, our research is a stepping stone on the road to the development of a system to identify and locate brick kilns for action against forced labour and environmental pollution.

VI. CONCLUSION

Automatic detection and localization of objects in satellite images have a great significance in the area of brick kilns. After extensive research and analysis, our study proposes the use of CNNs for the detection of kilns and the implementation of the YOLOv3 algorithm for the localization of kilns. The results of a 97.2% accuracy for detection combined with a mean IoU of 0.75, the precision of 0.950 and recall of 0.904 for localization showed how these algorithms could help develop an automatic solution to aid the government to take action against forced labour and environmental pollution. The results of our research have been promising thus far and can be further enhanced by adding more findings in this domain using the newer variants of YOLO introduced in the year 2020 and by implementing a GUI or creating an application to turn this into a usable software for detection and localization of kilns.

REFERENCES

[1] A. Eil, J. Li, P. Baral and E. Saikawa, "DIRTY STACKS, HIGH STAKES: An Overview of Brick Sector in South Asia," The World Bank, Washington, DC, 2020.

[2] N. DESK, "Pakistan third largest brick- producing country in South Asia," Business Recorder, 2017.

[3] D. S. Boyd, B. Jackson, J. Wardlaw, G. M. Foody, S. Marsh and K. Bales, "Slavery from Space: Demonstrating the role for satellite remote sensing to inform evidence-based action related to UN SDG number 8," ISPRS Journal of Photogrammetry and Remote Sensing, no. 142, pp. 380-388, 2018.

[4] U. Nazir, N. Khurshid, M. A. Bhimra and M. Taj, "Tiny-Inception-ResNet-v2: Using Deep Learning for Eliminating Bonded Labors of Brick Kilns in South Asia," arXiv:1907.05552, 2019.

[5] G. M. Foody, F. Ling, D. S. Boyd, X. LI and J. Wardlaw, "Earth observation and machine learning to meet sustainable development goal 8.7: Mapping sites associated with slavery from space,," Remote Sensing, vol. III, no. 11, p. 266, 2019.

[6] B. M. Skindar, A. Q. Sheikh, A. K. Pandit and B. A. Ganai, "Brick kiln emissions and its environmental impact: A Review.," Journal of Ecology and The Natural Environment, vol. I, no. 6, pp. 1-11, 2014.

[7] World Health Organization, [Online]. Available: https://www.euro.who.int/en/health-topics/noncommunicable-diseases/chronic-respiratory-diseases/data-and-statistics. [Accessed 10 August 2020].

[8] E. Krishnamoorthy, R. K. Sharma and T. Chithambarathanu, "Examining the variations of ground level ozone and nitrogen dioxide in a rural area influenced by brick kiln industries," Indian Journal of Science and Technology , vol. III, no. 8, 2010.

[9] I. Mukhtar, "With smog season looming, Pakistan shuts polluting brick kilns.," Reuters, Islamabad, 2018.

[10] M. Reynolds, "Volunteers teach AI to spot slavery sites from satellite images," NewScientist, 2017.

[11] T. Malisiewicz, A. Gupta and A. A. Efros, "Ensemble of Exemplar-SVMs for Object Detection and Beyond," in 2011 International Conference on Computer Vision, Barcelona, 2011.

[12] M. Xie, N. Jean, M. Burke, D. Lobell and S. Ermon, "Transfer Learning from Deep Features for Remote Sensing and Poverty Mapping," Stanford University, 2016.

[13] M. J. Khan, A. Yousaf, N. Javed, S. Nadeem and K. Khurshid, "Automatic Target Detection in Satellite Images using Deep Learning," Journal of Space Technology, vol. I, no. 7, 2017.

[14] W. Guo, W. Yang, H. Zhang and G. Hua, "Geospatial Object Detection in High Resolution Satellite Images Based on Multi-Scale Convolutional Neural Network," Remote Sensing, 2018.

[15] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," arXiv, 2016.

[16] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," arXiv, 2016.

[17] K. G. S. Dixit, M. G. Chadaga, S. S. Savalgimath, G. R. Rakshith and N. K. M R, "Evaluation and Evolution of Object Detection Techniques YOLO and R-CNN," International Journal of Recent Technology and Engineering, vol. 8, no. 2s3, 2019.

[18] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv, 2018.

[19] T.-Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollar, "Focal Loss for Dense Object Detection," arXiv, 2018.