Industrial Cognitive Ergonomics and Engineering Psychology, Vol. 35, 2022, 18–29
https://doi.org/10.54941/ahfe1001597

Image Classification for Project-based Learning to Differentiate Diagrams and Figures

Asmara Safdar[1,3], Sara Ali[1,2], Muhammad Sajid[1,2], Umer Asgher[1], and Yasar Ayaz[1]

[1]Human-Robot Interaction (HRI) Lab, Department of Robotics and Artificial Intelligence, School of Mechanical and Manufacturing Engineering (SMME), National University of Sciences and Technology (NUST), Islamabad, Pakistan
[2]School of Interdisciplinary Engineering & Science (SINES), NUST, Islamabad, Pakistan
[3]COMSATS University Islamabad Lahore Campus, 54000, Pakistan

ABSTRACT

This paper describes creation of a dataset and addresses an image processing problem in the field of education. A Convolutional Neural Network (CNN) based model is trained to classify the images extracted from academic documents. With the advent of distant learning mode and assessment criteria based on online submissions, there is a need to improve assessment approaches other than finding plagiarism. To enhance the understanding of the concepts, project-based learning (PjBL) in distant learning mode (DL) can be adopted. PjBL has proven successful even for complex engineering problems. It has been found out that PjBL of basic teaching assessment decreases the pressure on institutional resources while also making it easier and more practical for students. So, we are considering project reports or assignment as core source of evaluation. Extracting diagrams and software generated images (graphs and software generated object models) is focus for the current work as they reflect knowledge and main effort of a student especially in engineering academics. Here figures are refered as images of schematic representation to show the working or architecture of a work or a phenomenon. Software based images (sbi) include graphs, simulation images and software generated pictures or models. We aim to distinguish the *diagrams* and *sbi* from rest of the figures so it can be filtered out for further assessment. The data extracted is in the form of images. A CNN based classification model MobileNet is used to classify the images. The results show viability of the dataset and promising trend keeping in view the difficulty level of problem and size of dataset. Accuracy can be improved by adopting other approaches to train and clean data and also by increasing the data set by extracting more images from same domain of problem.

Keywords: Image classification, Artificial neural networks, Covolutional neural networks, Project based learning

INTRODUCTION

Project based learning is a method to make pupils use their acquired knowledge to address an enquiry question in the project (Bell, S. 2010). This has been applied and proved effective in terms of students' motivation to learn, improved confidence in self- directed learning activities, to enhance their problem-solving skills and working in a team as described by many researchers including (Ho-Quang, T., et al. 2014), (Ali, S. et al., 2020) and (Farid, T. et al., 2021). PjBL has also been proposed to adopt on undergraduate level especially for technical degrees. (Frank, M., Lavy, I. and Elata, D., 2003) present engineering students' and resource persons' views on PjBL as an effective learning environment for engineers.

Despite many advantages, PjBL poses many challenges in implementation and assesment because of its inherent complexities. A review study by (Guo, P at al., 2020) concludes that most of the work reported is in terms of effective outcomes based on self-reported measures that can be improved by introducing artifacts in the evaluation. (Williams, S., 2017) addresses problems faced during grading of the student projects and sheds light on different aspects of group assessment like co-created assignments, weitage and validation. One important aspect of grading a written assignment in this learning technique is checking the quality and similarity of the written contents among peers. (Chu, S.K.W. 2021) propose a model to prevent plagiarism in PjBL that encourages students to cite the work properly. It also points out posible reasons of plagiarism can be lack of understanding and technical or pedagogical issues. In other studies, (Adeva, J.J.G., et al., 2006) and (Pratama, H. et al., 2019) develop systems for plagiarism calculation for engineering documents for PjBL based on text.

However, grading a project document, particularly for engineering disciplines, merely based on text similarity is not sufficient to evaluate its quality; document images impart knowledge and effort exerted by a student, specifically PjBL. For this, (Meuschke, N. 2018), (Iwanowski, M. 2016), propose hybrid tools to detect both text and image plagiarism. In addition, to grade PjBL documents based on text and images, assessment and investigation of diagrams is very important. Many types of images can be added in academic documents such as graphs, software generated models and objects, diagrams drawn using basic shapes and computer aided tools. All these images have their own importance and contribution in an assignment. Sometimes, it is inevitable to filter the images based on their types for assessment purposes.

Applications of Artificial Intelligence (AI) based approaches have proven very useful in academia. In teaching industry, machine learning has been applied in various applications including prediction of future nature of teaching environments, plagiarism checking tools and softwares to aid learning for different padagogical techniques. For automated extraction of information from documents, many researchers have contributed their work such as (Wang, C. et al., 2018), (Paliwal, S.S. et al., 2019) and (Fan, J. et al., 2012). The extracted information can also be deployed to grade such documents based upon the category of images required for the project query. For example, for a case study of pipeline network, more diagrams are expected to be in the documents as compared to software-based images. Similarly, for an interdisciplinary study like integration of a machine learning algorithm for acquiring useful data insights, more graphs aka software-based images are added in the submissions. To reduce the grading time and speed up the plagiarism detection process (if employed) there is a need to extract these specific categories of images.

Current work is to address such problem where there is needed a tool to distinguish the images based on whether they are diagrams - hand sketched models, flow design of a procedure - or software-based images(sbi) - graphs, simulation models, software generated models. It involves the data extraction from projects and assignments submitted by engineering students of undergraduate level with implementation of PjBL. The dataset comprises of flow charts, diagrams, graphs, and images for classification generated from engineering assignments from over 100 students of graduation level. It is annotated manually in two classes: class1 is *sbi* and class2 is *figures*. There are approximately 1100 images of which 60% are instances of class1 and 40% are of class2. This is a supervised binary classification problem that is addressed by training (transfer learning) a Convolutional Neural Network (CNN). A Convolutional Neural Network is a deep neural network, which is designed to learn the latent and intrinsic features from 2D or 3D images. These features are appropriate in categorizing the images easily as diagrams involve more geometrical images. Sigmoid is used in the output layer for two mutually exclusive classes. The proposed model gives over 67% accuracy that is a promising value keeping in mind the limited dataset and potential to improvement by increasing training instances.

RELATED WORK

In research, numerous studies have done to evaluate the effectiveness and finding out ways to enhance difference aspects of PjBL pertaining to the study disciplines where it is being applied.

(Ali, S. et al., 2020) provide survey-based evidence of the usefulness of BjBL for students of complex engineering studies. (Ibrahim I. et al., 2018) propose one such approach in which they assigned a subject of Pipe Network Analysis for engineering students who applied their acquired knowledge efficiently. Similarly, (Farid, T. et al., 2021) implemented a robotics inspired transdisciplinary PjBL variant for undergraduate engineering students to confirm its sustainability.

(Karasneh, B. and Chaudron, M.R., 2013) study different machine learning approaches to extract UML class diagram images from documents. (Fu, L. and Kara, L.B., 2011) present a CNN based method to extract features from network like diagrams; hand sketched or made using computer aided design tools.

(Elyan, E. et al., 2018) generated a labeled dataset from real world engineering drawings and analyzed a few machine learning algorithms to detect and localize symbols within these drawings. (Kang, S.O. et al. 2019) use digital image processing techniques to extract design information and generate drawings from the data extracted from imaged piping and instrumentation diagram.

(Yun, D.Y. et al., 2020) implement an object-detection method to recognize graphical symbols in piping and instrument diagrams. They implemented unsupervised learning methods, k-means, and deep adaptive clustering and a convolutional neural network to perform this task.

Although there has been done a great deal of work in different directions of image detection and classification for engineering diagrams, there is a little clue to classify the said categories of images in engineering documents for PjBL.

METHODOLOGY

Dataset Generation

Dataset generation involves extracting all images from pdf and MS word documents of students' assignments and term-end projects. Initially, the data set consisted of over 1400 images in total with jpeg, png and jpg extensions. Each image was annotated manually by two annotators and validated by following the set of rules agreed upon by the annotators as given previously by (Hanbury, A, 2008). Inter-Annotator Agreement (IAA) was computed on the entire dataset. Both annotators agreed on 1103 images, disagreed on 297 (IAA = 93%), while 297 images were discarded by one or both annotators. This process posed many challenges keeping in mind the usefulness and mode of making images in the documents. Various images were composite having more than one type of image e.g., there were some figures with different images marking the levels and flow of procedures. Several images were composed of more than one image of same type those were split into multiple images.

The rules for labeling a *figure* were to include all images in this category that was any one of the following: 1) a diagram showing the flow of a procedure or architectural model either hand drawn or using online tools (as shown in Figure 1). 2) a model picture of an object, hand drawn or with the help of basic drawing tools (see Figure 2). 3) A model or diagram that includes software generated or a photographed image but there is a greater portion of labels and drawings (Figure 3). There was total 442 images in this class that constitutes 40% of total dataset.

To categorise images in sbi class, it should fall under either of these: 1) A graphical representation of any type of data [fig4], 2) An image produced through complex mathematical representations by a simulator [fig5], 3) Software generated models of real objects [fig6]. This class comprises 661 images that is 60% of total dataset.

Although each document had another class of images those were in small fraction proportional to other two categories, adding them to the dataset for another class wouldn't make a balanced dataset so they were discarded before further processing. Such images were mainly the photographs of real plants, models of engineering tools and machines and environment setups for experiments. These images have their own importance especially while depicting the details of such real-world setups. To include this category, either there was a need of large number of documents comparable to other two classes or data cleaning approaches to fit the minor class well would have to be adopted that is beyond the scope of current research. It is also important to mention that the dataset has different quality of images some of which are as small as few KB.

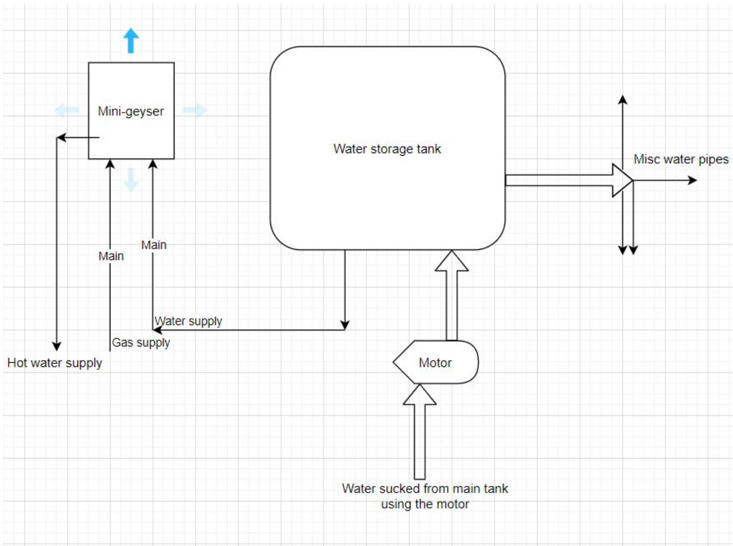

Figure 1: Basic line diagram of pipe-network as example of case 1 -images.

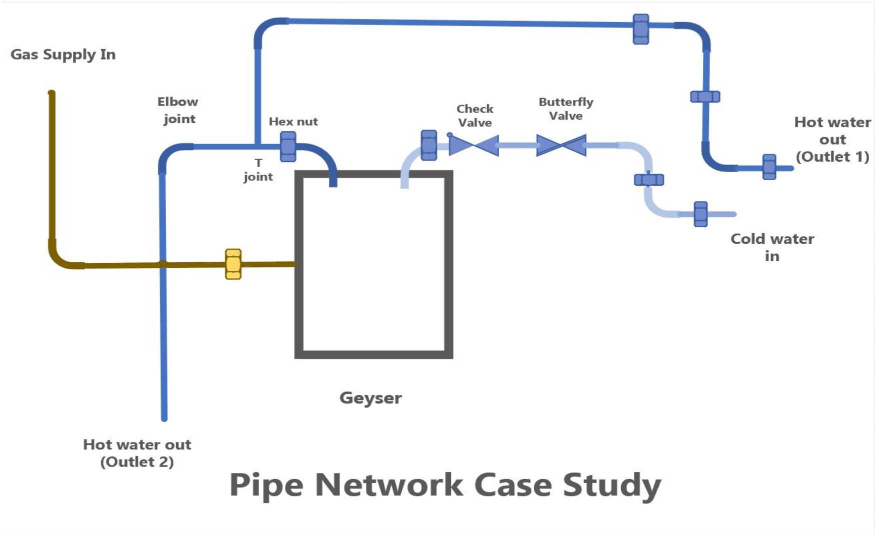

Figure 2: Detailed diagram with pipe-network components case 1-*images* example.

Model Selection and Training

The problem to address was to distinguish a figure image from a software-based image. Different convolutional neural network-based models can be employed for this classification, but we adopted MobileNet model because of its simplification and equally good accuracy comparable to other intricate classification models (Howard et al. 2017). It is accessible publicly in keras library of python.

Last three layers of the model were removed, and the trainable parameters were trained with our dataset. To classify the images, we used *sigmoid* activation function in the output layer while *adam* as the optimizer. Since the dataset is balanced i.e., 60%-40% instances for *sbi* and *figures* respectively,

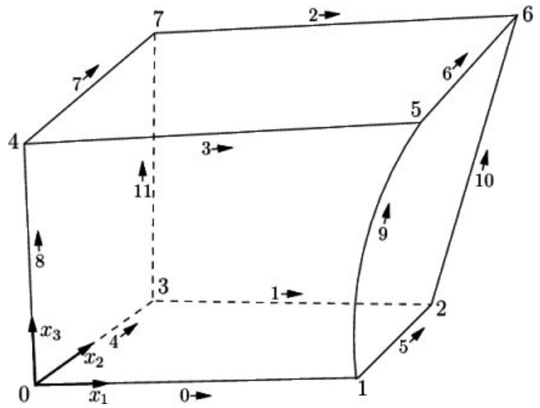

Figure 3: An example of figure images for case 2-*images*: A block structure of mesh element.

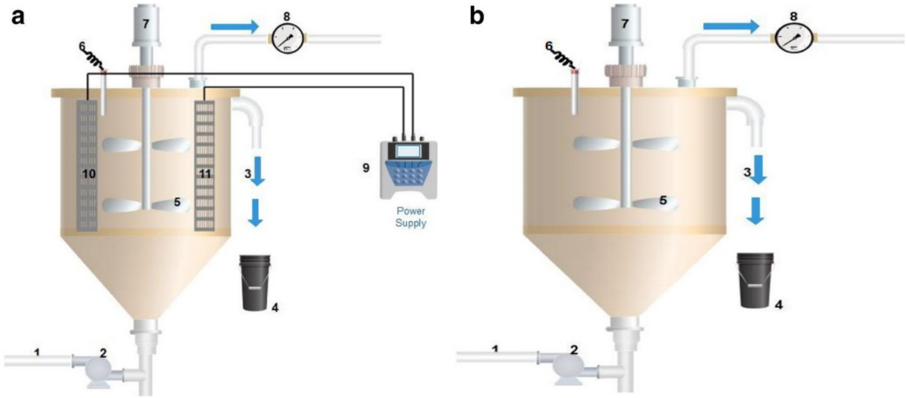

Figure 4: Figure image example for case 2-*images*: A CAD model of a plant machinary.

Table 1. Dataset statistics.

Classes	Training	Validation	Testing
figures	309	100	33
sbi	462	140	50

we have used *accuracy* as a performance measure to evaluate the model. We put 70% data into training set, and 30 % in validation and test set (Table 1). Each image was resized to 244, 244, 3 before feeding to network.

RESULTS AND DISCUSSIONS

MobileNet model gives an accuracy of 67% that shows a great potential to improve by increasing dataset, introducing appropriate methods for data cleaning and trying other types of CNNs. Precision and recall values for each class is 0.62 and 0.73 as shown in Table 1.

Data Cleaning of the current dataset can be done by many ways including denoising the minority class and enhancing the recognition rate as proposed

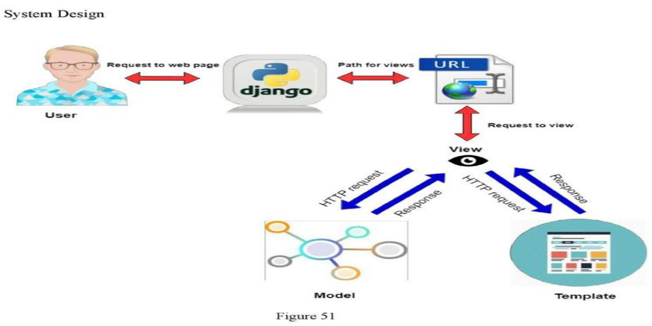

Figure 5: An example of case 3-*images*: Software architecture of a web application.

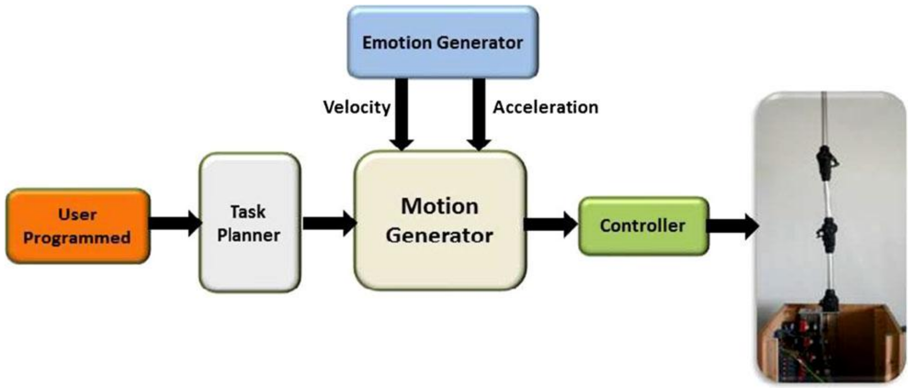

Figure 6: Flow diagram as examples of case 3-*images*: A part of architecture of a robot's actuator.

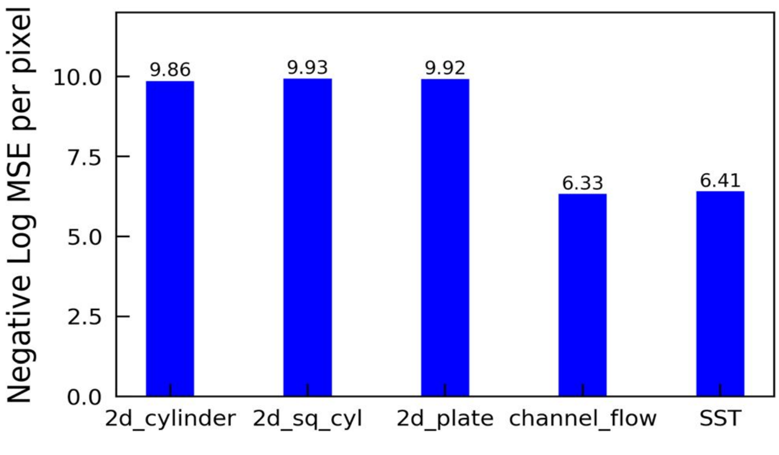

Figure 7: Bar graph for cgi case 1-*cbi*.

by (Zhang, Y. et al 2020), or based upon the features and nature of dataset as pointed out and proposed by (Corrales, D.C., et al. 2020)

The challenging aspect in classifying the images is due to the fact that since many images can be confused for both classes, like many images those are actually drawing figures of object models can be confused with sbi. Similarly,

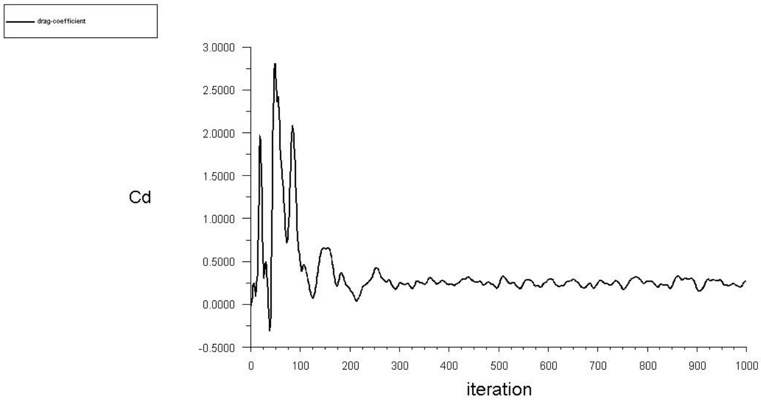

Figure 8: Aline graph for cgi case 1-*cbi*

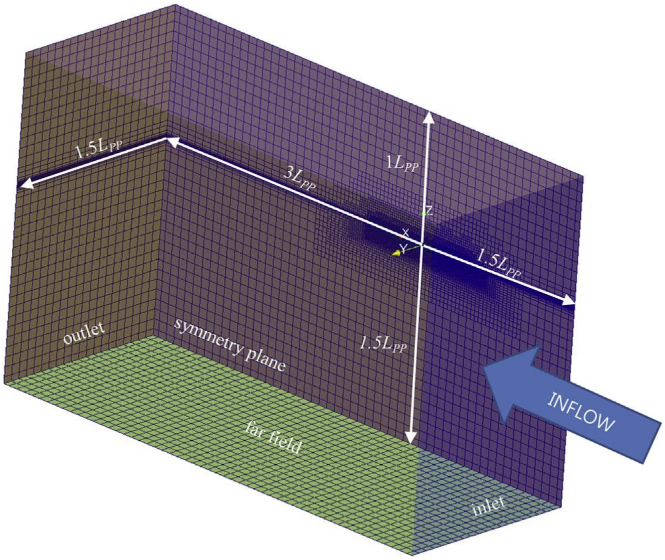

Figure 9: A block mesh of a cube for cgi case 2-*cbi*.

Table 2. Evaluation measures of the model.

Classes	Precision	Recall	Accuracy
sbi	62	73	67
figures	73	62	

many line graphs or bar graphs look like simple figures because of the basic lines and shapes in drawings. Moreover, all the figures are not simple with plain back grounds and simple basic shapes; some are the images of hand drawn figures and some are complex architectural diagrams with tangled shapes and various colors.

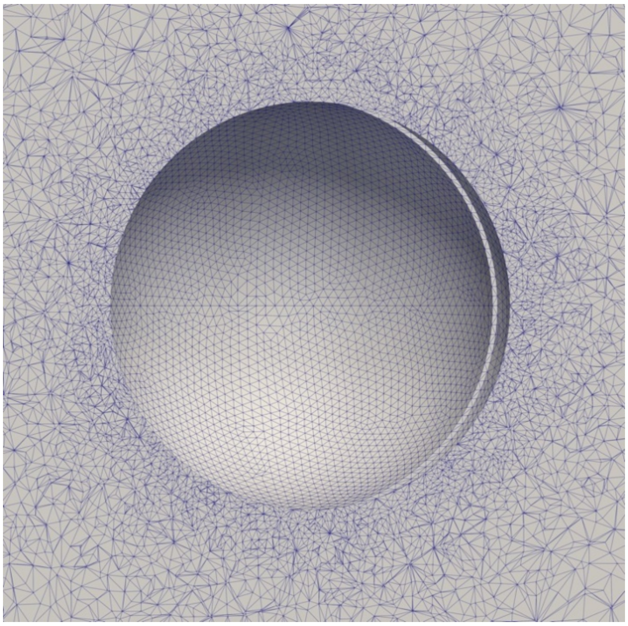

Figure 10: An unstructured mesh for a cricket ball for cgi case 2-*cbi*.

Figure 11: Case 3-*cbi* examples for software generated environment with different objects.

CONCLUSIONS

This paper presents the details of dataset generation for binary classification problem of images. The source of dataset is assignments and project reports of engineering students of undergraduate level. The purpose of creating such a dataset is to enhance the assessment criteria of project-based learning by focusing on main parts of a project document of technical disciplines particularly engineering. Since diagrams and software generated images are crucial for assessment, there should be a mechanism to check their similarity and

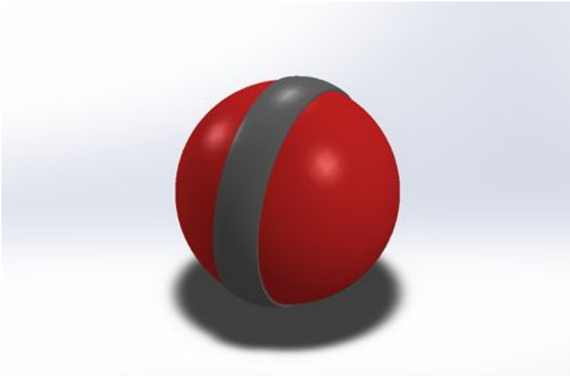

Figure 12: A cricket ball image generated by a simulation software as example of cgi case 3-*cbi*.

quality. The trained model presented in this work is helpful for classifying these two categories of images extracted from documents automatically.

FUTURE WORK AND APPLICATIONS

The current work is limited only to figures and software-based images and excludes all other types of images in test and training dataset. More accuracy can be gained by increasing the training examples, adopting data cleaning ways and trying other approaches in CNN. To increase scope of the present work, we can incorporate another category of images present in the types of documents under study i.e., photographed images of real architectures or object models. In addition, it can further be used to see the behavior of different models for classification of diagrams independent of the domain (in categories other than engineering). Also, it can be converted to train graph neural network (GNN) to see the potential of GNN for this type of classification.

ACKNOWLEDGMENT

The authors would like to acknowledge School of Mechanical and Manufacturing Engineering (SMME), National University of Sciences and Technology (NUST), Pakistan and European Union (EU)'s Horizon 2020, Research and Innovation Staff Exchange Evaluations (RISE) under grant agreement No 823904 - ENHANCE project (MSCA-RISE 823904) for technical support and funding.

REFERENCES

Adeva, J.J.G., Carroll, N.L. and Calvo, R.A., 2006, July. Applying plagiarism detection to engineering education. In *2006 7th International Conference on Information Technology Based Higher Education and Training* (pp. 722–731). IEEE.

Ali, S., Sajid, M., Ayaz, Y. and Asgher, U., 2020, July. A Methodology for Integrating Project Based Learning Outcomes and Attributes via Questionnaire. In *International Conference on Applied Human Factors and Ergonomics* (pp. 369–374). Springer, Cham.

Bell, S., 2010. Project-based learning for the 21st century: Skills for the future. The clearing house, 83(2), pp. 39–43.

Chu, S.K.W., Li, X. and Mok, S., 2021. UPCC: A model of plagiarism-free inquiry project-based learning. *Library & Information Science Research*, *43*(1), p.101073.

Corrales, D.C., Ledezma, A. and Corrales, J.C., 2020. A case-based reasoning system for recommendation of data cleaning algorithms in classification and regression tasks. Applied soft computing, 90, p.106180.

Elyan, E., Garcia, C.M. and Jayne, C., 2018, July. Symbols classification in engineering drawings. In 2018 International Joint Conference on Neural Networks (IJCNN) (pp. 1–8). IEEE.

Fan, J., Kalyanpur, A., Gondek, D.C. and Ferrucci, D.A., 2012. Automatic knowledge extraction from documents. IBM Journal of Research and Development, 56(3.4), pp. 5–1.

Farid, T., Ali, S., Sajid, M. and Akhtar, K., 2021. Sustainability of Project-Based Learning by Incorporating Transdisciplinary Design in Fabrication of Hydraulic Robot Arm. *Sustainability*, *13*(14), p.7949.

Frank, M., Lavy, I. and Elata, D., 2003. Implementing the project-based learning approach in an academic engineering course. *International Journal of Technology and Design Education*, *13*(3), pp. 273–288.

Fu, L. and Kara, L.B., 2011. From engineering diagrams to engineering models: Visual recognition and applications. Computer-Aided Design, 43(3), pp. 278–292.

G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. (2017) "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint arXiv:1704.04861, 2017.

Guo, P., Saab, N., Post, L.S. and Admiraal, W., 2020. A review of project-based learning in higher education: Student outcomes and measures. International Journal of Educational Research, 102, p.101586.

Hanbury, A., 2008. A survey of methods for image annotation. Journal of Visual Languages & Computing, 19(5), pp. 617–627.

Ibrahim, I., B. Khalid, M.W., Shoukat, G. and Sajid, M., 2018, July. An integrated approach for the teaching of pipe network analysis using project based learning. In *Fluids Engineering Division Summer Meeting* (Vol. 51555, p. V001T04A003). American Society of Mechanical Engineers.

Iwanowski, M., Cacko, A. and Sarwas, G., 2016, September. Comparing images for document plagiarism detection. In *International Conference on Computer Vision and Graphics*(pp. 532–543). Springer, Cham.

Kang, S.O., Lee, E.B. and Baek, H.K., 2019. A digitization and conversion tool for imaged drawings to intelligent piping and instrumentation diagrams (P&ID). *Energies*, *12*(13), p.2593.

Karasneh, B. and Chaudron, M.R., 2013, March. Extracting UML models from images. In 2013 5th International Conference on Computer Science and Information Technology(pp. 169–178). IEEE.

Meuschke, N., Gondek, C., Seebacher, D., Breitinger, C., Keim, D. and Gipp, B., 2018, May. An adaptive image-based plagiarism detection approach. In *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries* (pp. 131–140).

Meuschke, N., Stange, V., Schubotz, M. and Gipp, B., 2018, June. HyPlag: a hybrid approach to academic plagiarism detection. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval* (pp. 1321–1324).

Paliwal, S.S., Vishwanath, D., Rahul, R., Sharma, M. and Vig, L., 2019, September. Tablenet: Deep learning model for end-to-end table detection and tabular data extraction from scanned document images. In *2019 International Conference on Document Analysis and Recognition (ICDAR)* (pp. 128–133). IEEE.

Pratama, H. and Prastyaningrum, I., 2019, February. Effectiveness of the use of Integrated Project Based Learning model, Telegram messenger, and plagiarism checker on learning outcomes. In *Journal of Physics: Conference Series* (Vol. 1171, No. 1, p. 012033). IOP Publishing.

Wang, C., Ma, X., Chen, J. and Chen, J., 2018. Information extraction and knowledge graph construction from geoscience literature. *Computers & geosciences*, *112*, pp. 112–120.

Williams, S., 2017. Investigating the allocation and corroboration of individual grades for project-based learning. *Studies in Educational Evaluation*, *53*, pp. 1–9.

Yun, D.Y., Seo, S.K., Zahid, U. and Lee, C.J., 2020. Deep neural network for automatic image recognition of engineering diagrams. Applied Sciences, 10(11), p.4005.

Zhang, Y., Jin, Z., Liu, F., Zhu, W., Mu, W. and Wang, W., 2020, March. ImageDC: Image Data Cleaning Framework Based on Deep Learning. In 2020 IEEE International Conference on Artificial Intelligence and Information Systems (ICAIIS) (pp. 748–752). IEEE.