Title of your Report

Given Name Surname
Faculty of Computer Science & Engg.
AI Research Group
GIK Institute of Engg. Sciences & Tech.
Topi, Khyber Pakhtunkhwa, Pakistan.
email address or ORCID

Given Name Surname
Faculty of Computer Science & Engg.
GIK Institute of Engg. Sciences & Tech.
Topi, Khyber Pakhtunkhwa, Pakistan.
email address or ORCID

***Abstract*—*CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Report Title or Abstract.**

Two sentences of background and importance of this topic.

Two sentences defining the gap analysis (work not done in field) and your problem statement.

Two sentences on how you performed your experimentation and the overall methodology deployed in this study.

Two sentences of Results and your findings.

One sentence for significance of the results and your contribution.

In total, this section must consist of 180 words.

***Index Terms*—keyword1, keyword2, keyword3, keyword4**

I. General instructions

Comment out this section after completing your write up.

1) You must have exactly the mentioned sections with the mentioned number of paragraphs.
2) DO NOT DELETE ANY TEXT FROM THIS REPORT. JUST COMMENT OUT AND WRITE THE RELEVANT TEXT UNDER EACH COMMENT.
3) Observe the report page limits (minimum 5 pages without references) and limit on maximum text is 7 pages without references.
4) In case of figures, keep your raw data table also stored as excel file in this repository.
5) Always draw your figures in powerpoint/draw.io/canva and then save it in pdf. In case of results from your code, directly save the images in pdf from your program. Avoid screenshots but in worst cases, where taking a screenshot is necessary, always zoom in to the max before taking the screenshot.
6) Each paragraph must consist of 7-10 sentences.
7) Add references where required.
8) Write one sentence (ending with .) per line of Latex.
9) Total references should be between 20 and 30.
10) Generate tables in Excel first, then convert them into latex tables by using https://www.tablesgenerator.com website and pasting the table into the site alongwith drawing its borders. Copy and paste this table into latex.
11) Use latest references with 80% references later than 2019.
12) Use Google scholar for finding references and not Google.

II. Introduction

One paragraph on introducing the field and the topic of interest.

One paragraph on the importance of the selected topic and why it is significant to work on it in recent times.

A. Related Work

One paragraph defining the work that has been done in this field, with a table summarizing the work that has been done in literature as shown below in Table I.

B. Gap Analysis

One paragraph defining what has not been done or what is still missing in the field (gap analysis).

C. Problem Statement

Following are the main questions addressed in this study.

1) Research Question 1.
2) Research Question 2.
3) Research Question 3.

D. Novelty of our work and Our Contributions

One paragraph explaining your approach and novelty/contributions of your work.

One paragraph on what you are doing in this report (your contributions) and a small one-two liner summary of your results.

III. Methodology

A. Dataset

One paragraph and one figure representing your dataset, also give references (citation) from where the dataset is available, and the labels/ground truth as shown in Figure 1 OR as in Figure 2.

B. Overall Workflow

One paragraphs defining your methodology through a flow diagram of your work as shown in Figure 4 OR in Figure 3.

C. Experimental Settings

(Optional) One paragraph for hyper-parameter settings and network architecture as shown in Table II.

(Optional) One paragraph for experimental settings of your and competing methods (if any).

TABLE I
LITERATURE REVIEW TABLE SHOWING THE CONTRIBUTIONS OF VARIOUS AUTHORS FOR QUANTIZATION OF NETWORKS.

Paper Name	FCNs Used	L2 Error Minim.	Applied on				Signal Quantized	Dataset used	No. of bits	Layerwise sens. Analysis	Sem. segm.
			Conv. Layer	Skip Layer	Trans. Layer	Fully Conn. Layer					
Vanhoucke et al. [1]						✓	✓	×			
Courbariaux et al. [2]			✓			✓		MNIST, SVHN, CIFAR-10	10		
Gupta et al. [3]			✓			✓		MNIST, CIFAR-10	12		
Proposed Approach	✓	✓	✓	✓	✓	×	×	Pascal VOC 2012	2,3,4,5	✓	✓

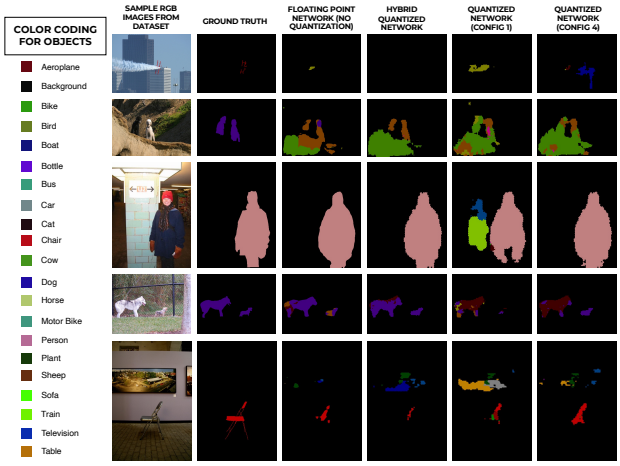

Fig. 1. Image showing some sample images present in the dataset, their pixel-wise labels and resulting pixel labels from floating point network, hybrid quantized network, and two configurations of quantized networks. The legend displays the color and class (name) of the object to be identified in the image. Five sample images containing aeroplane, dogs, person, and chair are shown along with their classification. The data and the pixel labels (ground truth) are taken from Pascal VOC 2012 dataset.

TABLE II
CONFIGURATION TABLE SHOWING THE NETWORK CONFIGURATION OF FCN USED IN THIS STUDY. THE TABLE SHOWS THE VARIOUS CONFIGURATION SETTINGS USED FOR FCN8.

Network Configuration	
Epochs	50
Learning rate	0.0001
Mini batch size	20
Optimizer	SGD
Momentum	0.9
Weight decay	0.0002
L_2 Regularization	None
Samples in training set	8498
Samples in validation set	786

IV. RESULTS

Three (or more) paragraphs explaining your results. At least one paragraph targeting one research question with at least one figure (preferably) or table (where figure is not possible). This section must contain only results and nothing else (not your own opinion or any sort of discussion on quality of results).

A sample figure is shown in Figure 5.

V. DISCUSSION

Three to four paragraphs discussing the results (at least one paragraph for each research question). Your opinion on how good/bad the results are. Draw inferences from the results here. Explain novelty of your contributions and what was missing that you have explored here. Any other point you would like to discuss related to this study.

A. Future Directions

One paragraph for what are the future directions in your opinion for continuing this study.

VI. CONCLUSION

One paragraph related to conclusions drawn from your whole experimentation.

In total this section must consist of 240-260 words.

REFERENCES

[1] V. Vanhoucke, A. Senior, and M. Z. Mao, "Improving the speed of neural networks on cpus," in *Deep Learning and Unsupervised Feature Learning Workshop, NIPS 2011*, 2011.

[2] J. David, M. Courbariaux, and Y. Bengio, "Training deep neural networks with low precision multiplications," *Computer Science*, 2014.

[3] S. Gupta, A. Agrawal, K. Gopalakrishnan, and P. Narayanan, "Deep learning with limited numerical precision," in *International Conference on Machine Learning*, 2015, pp. 1737–1746.

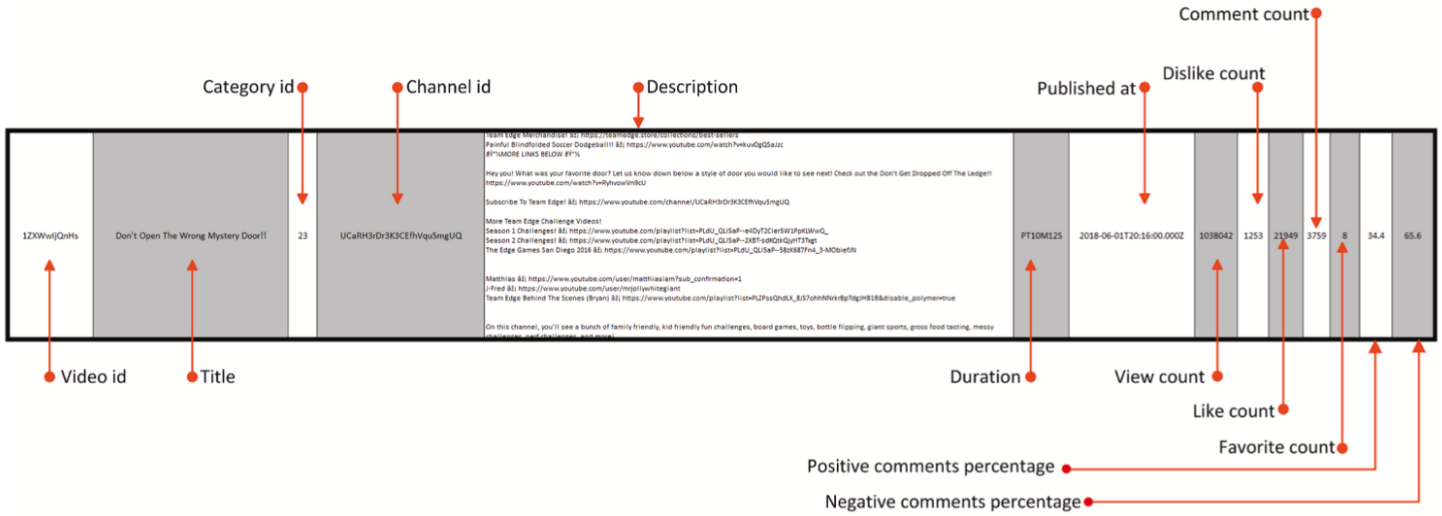

Fig. 2. A screenshot of the captured data.

Fig. 2. Image showing some sample images present in the dataset, their pixel-wise labels and resulting pixel labels from floating point network, hybrid quantized network, and two configurations of quantized networks. The legend displays the color and class (name) of the object to be identified in the image. Five sample images containing aeroplane, dogs, person, and chair are shown along with their classification. The data and the pixel labels (ground truth) are taken from Pascal VOC 2012 dataset.

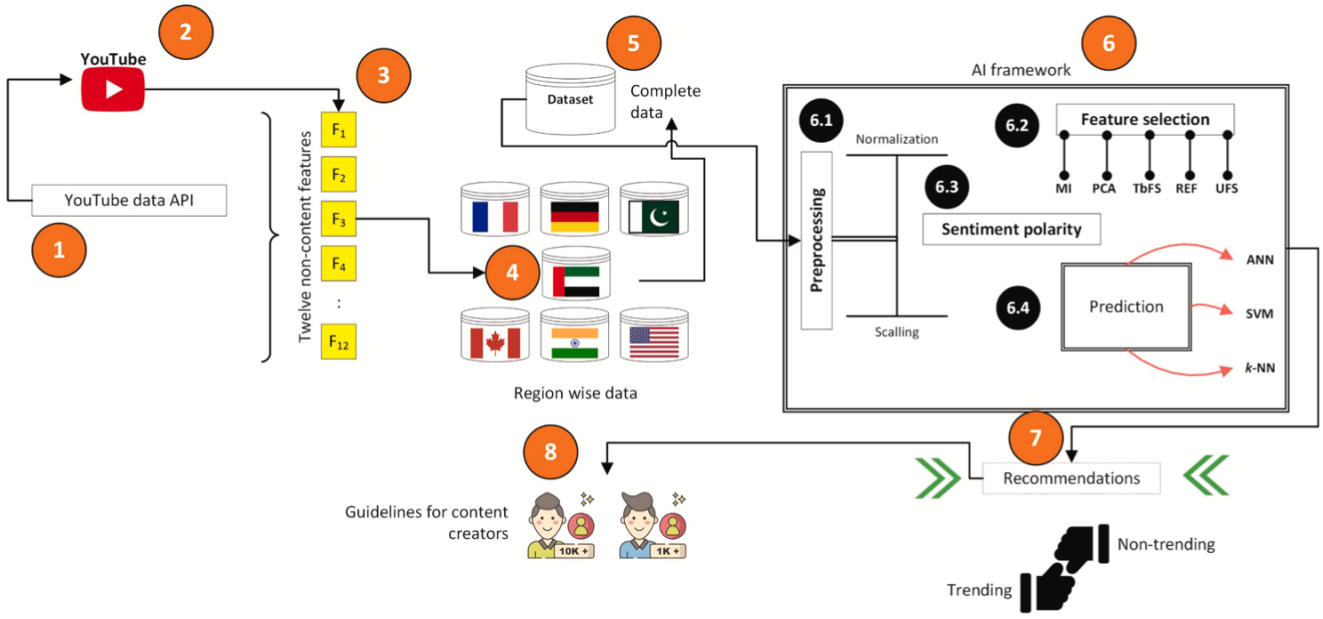

Fig. 1. The overall working of the proposed solution.

Fig. 3. Figure showing the flowchart proposed for FCN-8 quantization and the comparison pipeline followed (for quantization technqiues, i.e., Direct Quantization, Llyod's Quantizer and L_2 error minimization) in the current study based on pixel accuracy, mean IOU, and mean accuracy.

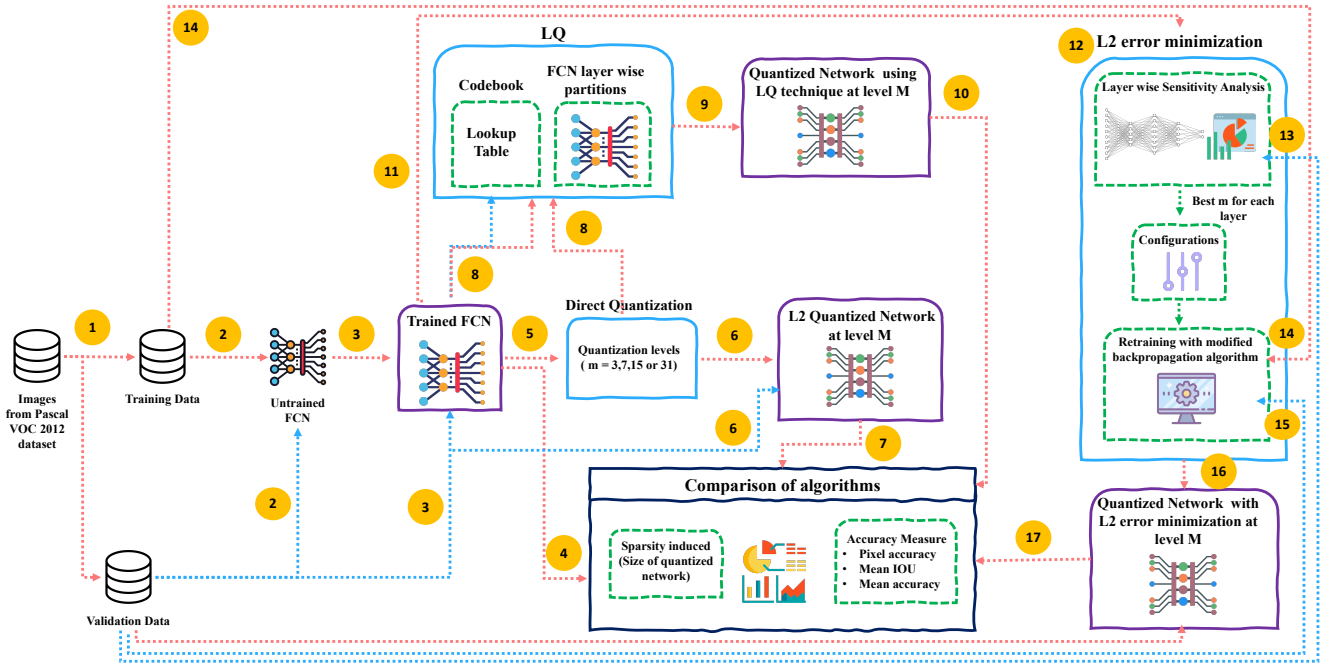

Fig. 4. Figure showing the flowchart proposed for FCN-8 quantization and the comparison pipeline followed (for quantization techniques, i.e., Direct Quantization, Llyod's Quantizer and L_2 error minimization) in the current study based on pixel accuracy, mean IOU, and mean accuracy.

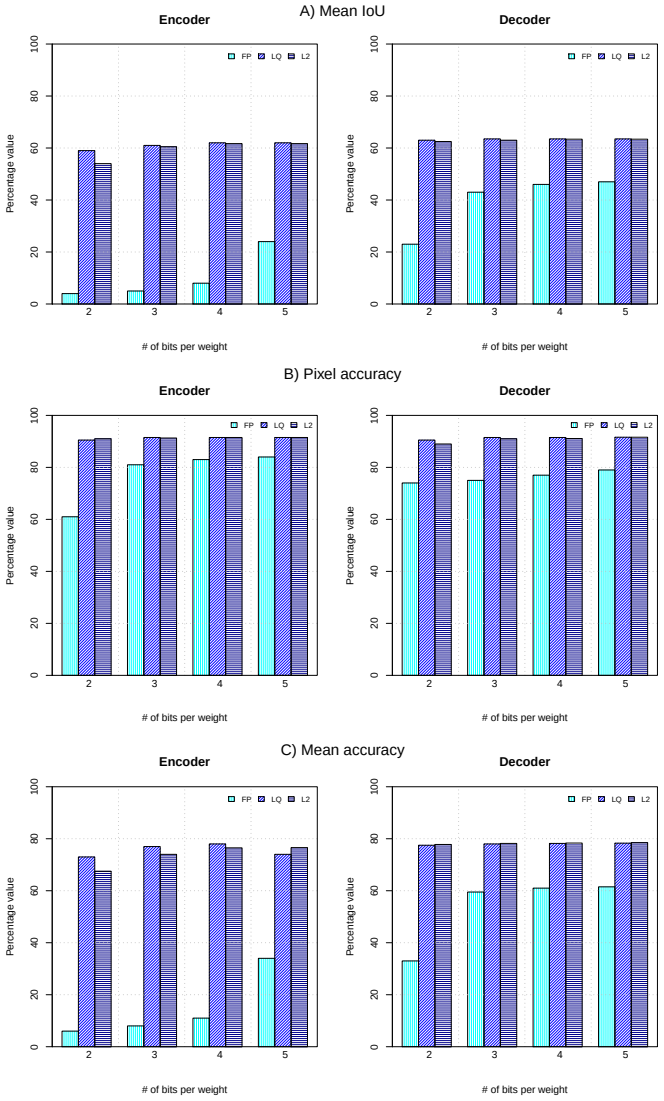

Fig. 5. Figure comparing the three quantization techniques Fixed Point (FP), Lloyd's quantizer (LQ) and L_2 error minimization (L_2) on the three performance metrics divided into encoder and decoder layers. Mean IoU is shown for the three techniques in Panel A), pixel accuracy in Panel B), and mean accuracy in Panel C) respectively. Note that FP is consistently worse than both LQ and L_2, while L_2 and LQ are of comparable accuracy. Also, FP is most sensitive to number of bits in all metrics while L_2 and LQ are relatively insensitive.