HH-CompWordNet: Holistic Handwritten Word Recognition in the Compressed Domain

[Paper Id:- 190]

Data Compression Conference
(Virtual)
March 2021

Bulla Rajesh[1], Priyanshu Jain[1], Mohammed Javed[1] and David Doermann[2]

Email: rsi2018007@iiita.ac.in

[1]*Department of IT*
Indian Institute of Information Technology Allahabad, U.P, India

University at Buffalo
The State University of New York

[2]*Department of CSE,*
University at Buffalo, NY, USA

Motivation

- Handwritten documents are common in historical and current daily communications.
- In general, handwritten recognition techniques suffer from many challenges, including uneven spacing between text lines and words, stroke curvature Variation, overlapping of adjacent text components and the presence of diacritics.
- The Better recognition of handwritten text open many applications in the field of Document Image Analysis (DIA)[1].

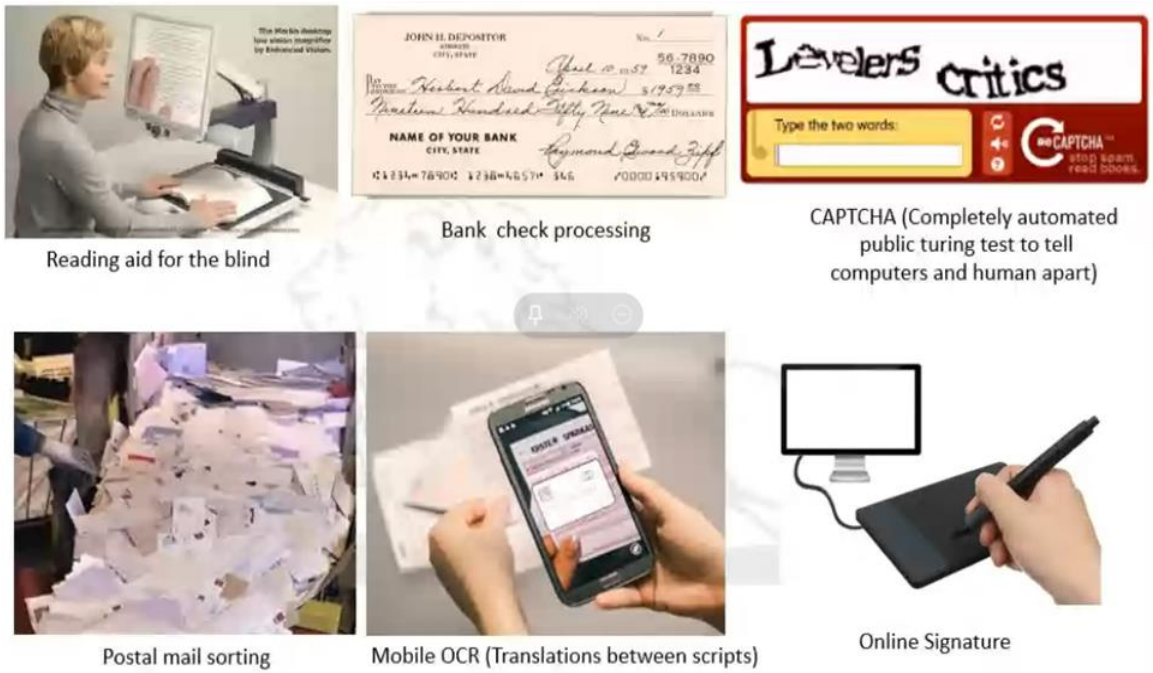


Fig 1. Various applications of handwritten text recognition.

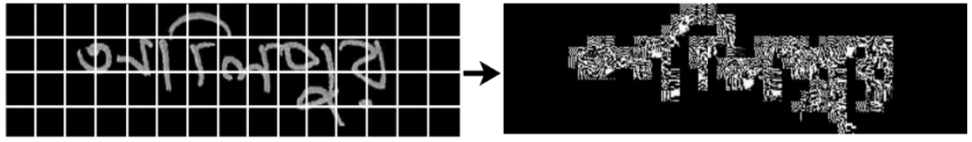
[1] M Javed, P. Nagabhushan, B.B. Chaudhuri, "A review on document image analysis techniques directly in the compressed domain", Artificial Intelligence Review, Volume 50, 539–568, 2018

OUTLINE

- INTRODUCTION
- METHODOLOGY
- PROPOSED DEEP LEARNING ARCHITECTURE
- EXPERIMENTS AND ANALYSIS
- CONCLUSION AND FUTURE WORK

Introduction

- In general handwritten words are recognized using two techniques: First one is Character based approaches and second one is Holistic based approaches.
- Character-based approaches follow a two-stage process where the individual characters are first identified and used to recognize the entire word.
- This approach needs domain expertise and includes more computational charges.
- Our paper is focused on developing holistic word recognition using a deep learning model directly in the **JPEG compressed domain**.



(a) (b)

Fig. 2 The encoding of (a) handwritten word image during 8×8 DCT transformation, (b) DCT encoded word image.

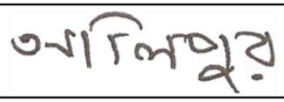
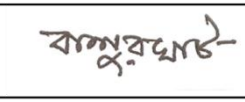
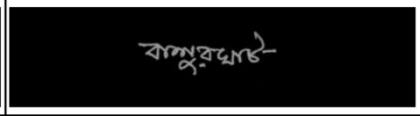
Word Name	Alipore	Balurghat
Original Images		
Pre Processed Image 100×400		
DCT Image 100×400		

Fig. 3 Sample Bangla handwritten word images in both pixel and DCT compressed representations [2]

- JPEG is one of the most used (More than 90% and 10:1 compression ratio) digital compression algorithms and became the default image compression format to many devices like mobile phones, digital cameras and so on.
- The frequent usage of these devices in a day to day life have resulted the large number of handwritten JPEG compressed document images. It compress the image by dividing it into 8×8 blocks and transforming the each using Discrete Cosine transformation(DCT).

[2] D. Das, D. R. Nayak, R. Dash, B. Majhi, and Y. Zhang, "**H-wordnet: a holisticconvolutional neural network approach for handwritten word recognition**", IET Image Processing, vol. 14, pp. 1794–1805, 2020.

Methodology

- Generation of DCT compressed word images and feeding them to HH-CompWordNet
- Recognition of handwritten words in compressed domain.

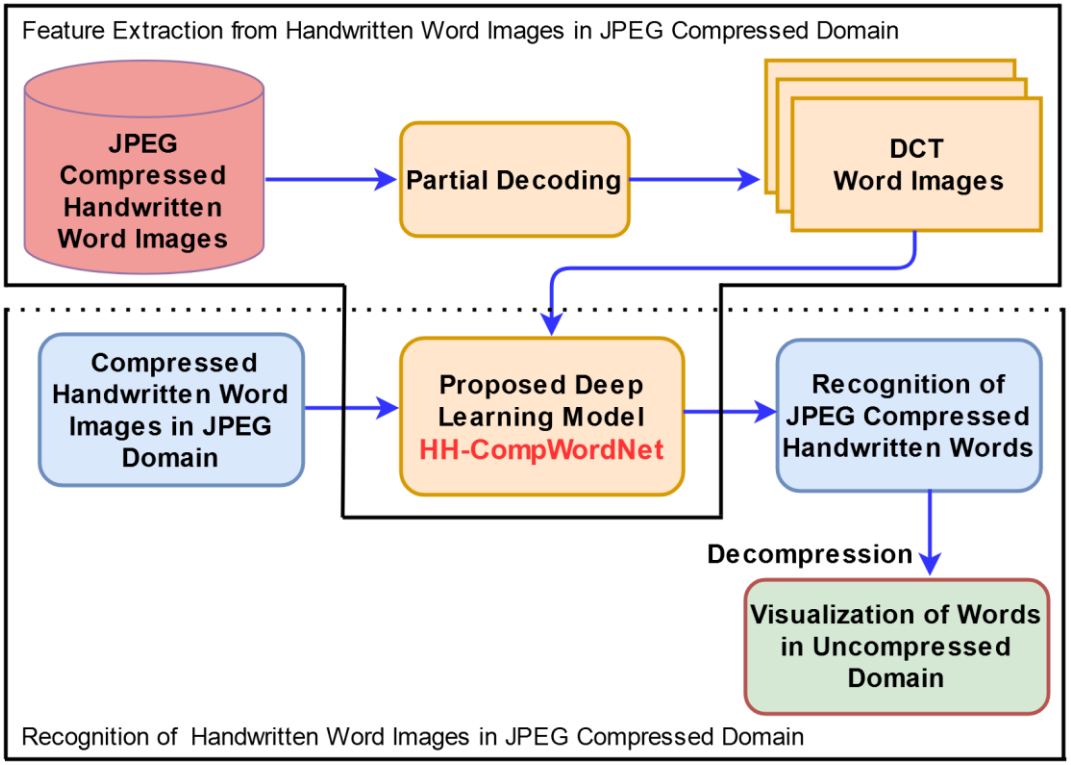


Fig.4 Block diagram of the proposed holistic compressed word recognition model in JPEG compressed domain.

Proposed Deep Learning Architecture (HH-CompWordNet)

- Feature Extraction from the DCT compressed word images using 3 convolution layers.
- Feeding the features to Fully connected layer for recognizing the compressed word.

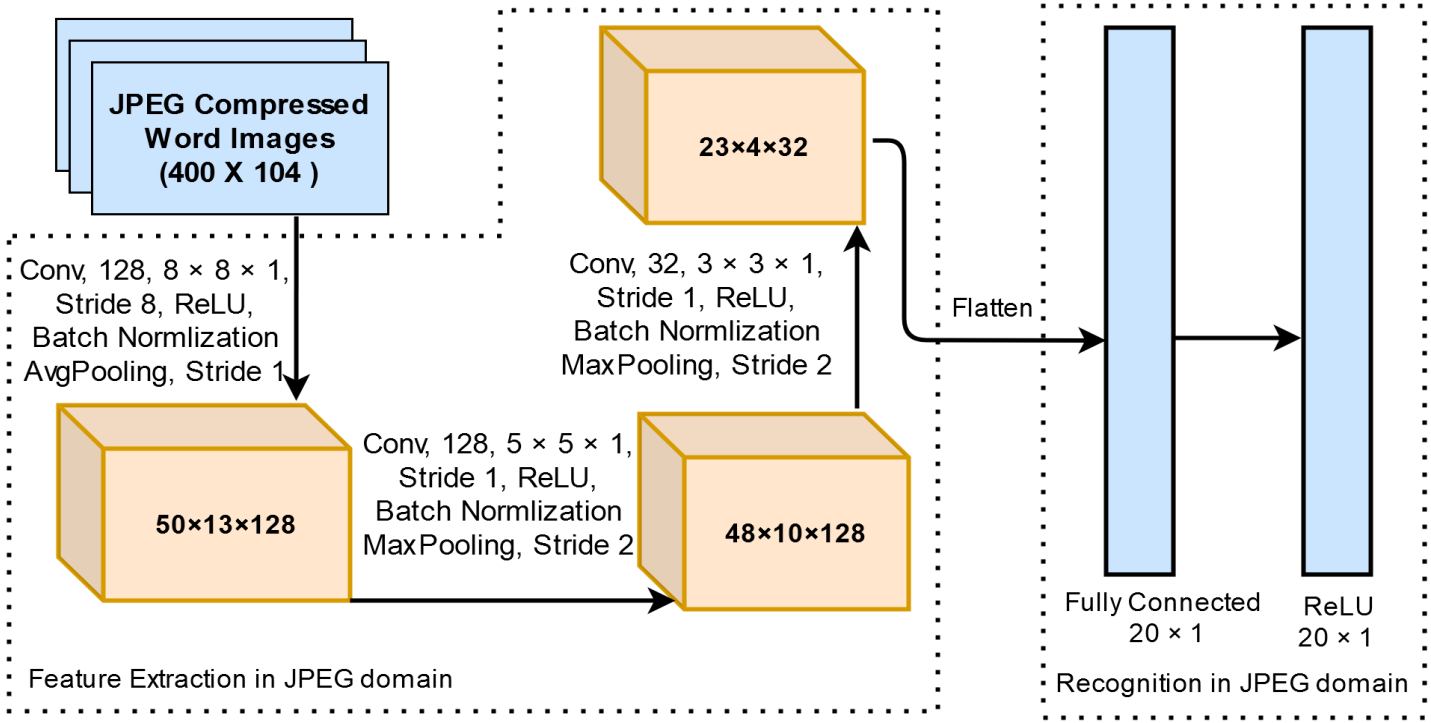


Fig. 5 Proposed deep learning architecture - HH-CompWordNet for the recognition of JPEG compressed word images

Visualization of Layer wise activations

- Handwritten feature activations in the compressed domain at various layers in the model.
- Encountered vanishing gradients problem as deep model processing text contents using DCT values.

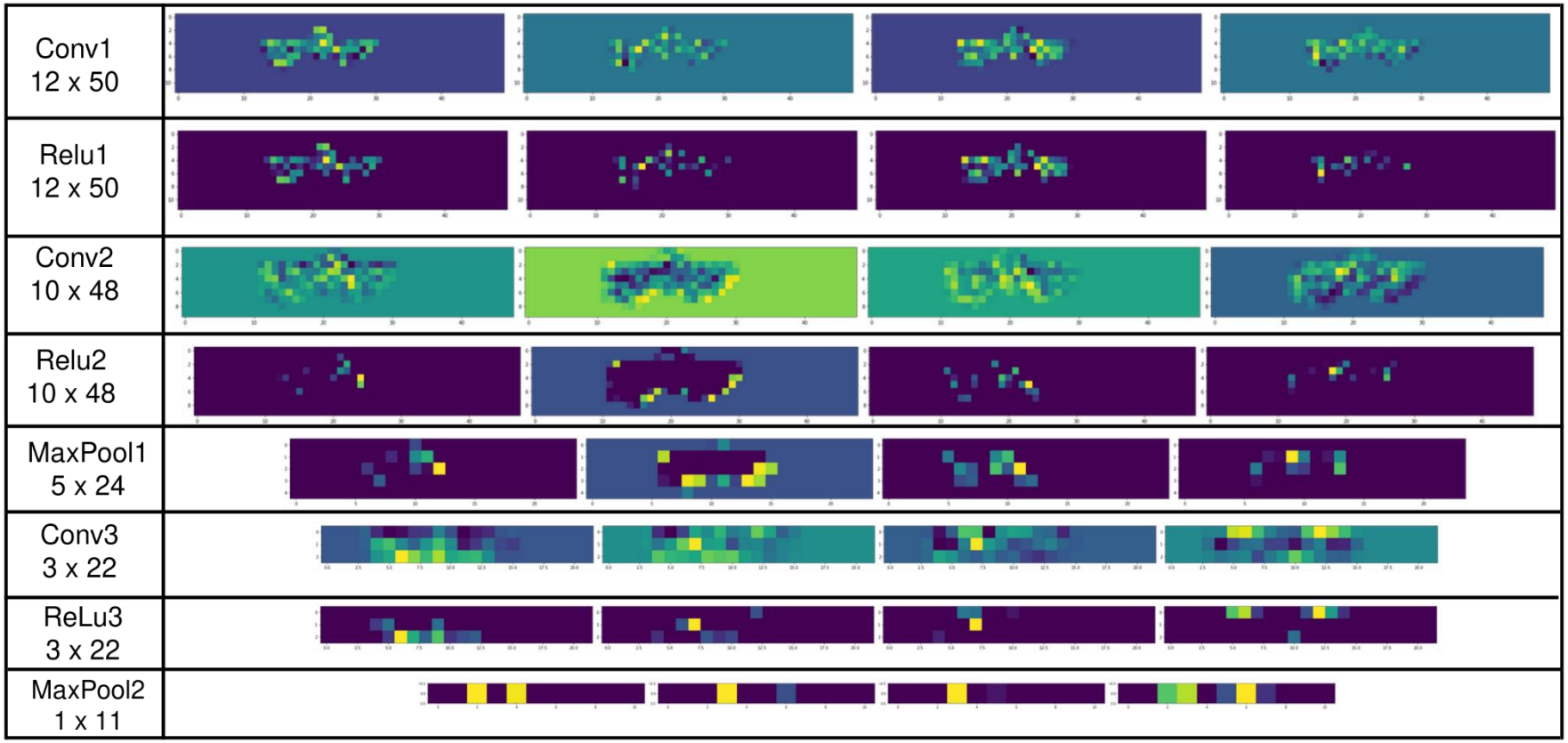


Fig. 6 The visualization of output activation at each convolution layer for the HH-CompWordNet

Experiments and analysis

Table 1: Performance of the proposed deep learning architecture - HH-CompWordNet on the JPEG compressed version of CMATERdb2.1.2.1 dataset compared to existing related uncompressed domain models.

Reference	Input Data Type	Dataset Used	Accuracy(%)
Bhowmik[12]	Pixel Data	CMATERdb2.1.2	83.64
Dibyasundar[3]	Pixel data	CMATERdb2.1.2	96.17
Proposed HH-CompWordNet	**JPEG Compressed Stream**	CMATERdb2.1.2.1	**86.80**

Table 2: Efficiency comparison of the proposed deep learning architecture - HH-CompWordNet on the JPEG compressed version of CMATERdb2.1.2 dataset in relation to existing deep learning models.

Reference	Number layers	Total Parameters	Train Time(in sec)	Accuracy (%)
Dibyasundar[3]	4	7,97,912	2510	96.17
ulicny[19]	4	7,50,000	7,200	86.35
Pistino[21]	-	-	-	82.00
Gueguen[20]	50	23M	10,800	-
Proposed HH-CompWordNet	**3**	**1,33,652**	**7,800**	**86.80**

Conclusion and Future work

Table 3: Accuracy for individual classes with the proposed model HH-CompWordNet tested on CMATERdb2.1.2 dataset.

Class	0	1	2	3	4	5	6	7	8	9
Precision	0.77	0.88	0.88	0.84	0.89	0.77	0.88	0.88	0.94	0.93
Recall	0.81	0.88	0.79	0.83	0.95	0.83	0.94	0.69	0.91	0.96
F1 Score	0.79	0.88	0.84	0.83	0.92	0.80	0.91	0.77	0.93	0.95
Support	54	49	58	52	41	48	48	51	55	56
Class	10	11	12	13	14	15	16	17	18	19
Precision	0.93	0.95	0.91	0.86	0.88	0.76	0.88	0.90	0.85	0.91
Recall	0.86	0.85	0.80	0.89	0.98	0.88	0.86	0.91	0.97	0.84
F1 Score	0.89	0.90	0.85	0.88	0.92	0.82	0.87	0.90	0.91	0.87
Support	49	41	49	47	43	51	42	57	60	49

- Handwritten word recognition in the JEPG compressed domain is explained with state-of-the-art recognition accuracy.
- We are motivated to continue in this domain to explore various applications directly in it.

Thank You

HH-CompWordNet: Holistic Handwritten Word Recognition in the Compressed Domain

[Paper Id:- 190]

Data Compression Conference
(Virtual)
March 2021

Bulla Rajesh[1], Priyanshu Jain[1], Mohammed Javed[1] and David Doermann[2]

Email: rsi2018007@iiita.ac.in

[1]Department of IT
Indian Institute of Information Technology Allahabad, U.P, India

University at Buffalo
The State University of New York

[2]Department of CSE,
University at Buffalo, NY, USA