IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 16 March 2023, accepted 25 March 2023, date of publication 5 April 2023, date of current version 13 April 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3264605

RESEARCH ARTICLE

Compressive Wavelet Domain Deep CNN for Image Classification Using Genetic Algorithm Based Sensing Mask Learning

BABA FAKRUDDIN ALI B. H. AND **PRAKASH RAMACHANDRAN**

School of Electronics Engineering, Vellore Institute of Technology, Vellore 632014, India

Corresponding author: Prakash Ramachandran (prakash.r@vit.ac.in)

ABSTRACT Using a novel Genetic Algorithm-based Compressive Learning (GACL), a compressed domain-learning framework is proposed that is implemented on the Haar wavelet approximation coefficient images of the standard kaggle RGB cat dog dataset with every images resized to 256 × 256 × 3. The compressive sensing (CS) measurements on the selected dataset is achieved by using a simple reduced pixel scheme by retaining only P% of the pixels of the approximation coefficient images and forcing the remaining pixels to 0 using the Primitive Walsh Hadamard (PWH) binary mask and the masked images are used for further learning. A numerical experiment is conducted to analyze the image classification performance of deep convolution neural network (DCNN) learning on compressive sensing (CS) measurements of wavelet approximation coefficient image of the selected dataset. The unmasked wavelet approximation coefficients images are of size only one fourth of the original image, but they visually resembles the original image. The numerical experiment shows that when learning is done on this unmasked wavelet approximation coefficient images a training accuracy of 97% and validation accuracy of 77% are achieved which is remarkable and as good as using the complete spatial domain image. It is found from numerical experiment that, when PWH masking with $P = 10$ is applied only to the test images the validation accuracy falls up to 58% and this fall is due to the fact that the training is done on unmasked images and tested on masked images. On the other hand in a compressive learning framework the DCNN is trained using masked images and when tested using masked images the validation accuracy rises up to 62% due to the fact both the trained images and test images are masked. Further, it is demonstrated that the best PWH masks can be learned by GACL in which the training accuracy increases up to 89% when vertical cross over is used in GACL and increases up to 96% when diagonal crossover is used in GACL. in this case of GACL using diagonal crossover, the 96% of training accuracy using only (10/4 = 2.5) 2.5% of the image is very remarkable and stand as proof of concept to implement compressive learning framework which need very less pixel thus very less measurement rate (MR) both in training and testing phase minimizing the bandwidth and storing requirement in applications related to IoT and cloud solutions. The average SSIM (S_{avg}) and average PSNR ($PSNR_{avg}$) are used as quality measurements, which reduce as P reduces, and it is demonstrated that the average SSIM and average PSNR improves when GACL is used to learn the mask, which is the key reason for the performance improvement.

INDEX TERMS Binary masking, compression sensing (CS), deep convolution neural network (DCNN), measurement ratio (MR), Haar wavelet, genetic algorithm-based compressive learning (GACL), Primitive Walsh Hadamard (PWH).

The associate editor coordinating the review of this manuscript and approving it for publication was Sangsoon Lim.

I. INTRODUCTION

Deep convolution neural network (DCNN) trained with multiple classes of images is a robust method to classify an image under the test as a subset of the trained classes. In general, the images used for training and testing are raw and uncompressed. When the test image is compressed, the models validation accuracy deteriorates. This can be overcome by training the models in compressed domain in which the compressed information are used directly for training and testing and many work are done to develop learning system based on compressed information [1], [2], [3], [4], [5], [6], [7], [8]. This enables faster communication and eliminates the need for reconstruction algorithm. Compressed domain inference is proposed by Calderbank et al. [9] in which they use support vector machine (SVM) learning on compressed data. Prof. Pavan Turaga and his team proved compressed domain classification [4] using the rate-adaptive network and Davenport et al. [5], used a smashed filter based on a random sensing matrix and a neighborhood classifier. Khmag Asem [10], [11] used different techniques for noise removal.

CS is a paradigm shift in the sampling and compression process, in which only the required number of measurements are done on the signal, the original signal is reconstructed back using optimization techniques [9], [10], [11], and the method of learning on CS measurements will enjoy the compactness of the CS [12], [13], [14], [15], [16], [17], [18], [19], [20].

Xu et al [19] developed a single neural network and used it to categorize data in any measurement rate (MR) using single-pixel camera (SPC) based CS approach as shown in Figure 1.

In this approach, the training set is a super set of data with multiple MR and thus the model is suitable for any MR. In this approach the images are subjected to CS measurements using SPC sensing matrix and the CS measurements are used in the training a CNN model. CNN model is robust machine learning framework that works very well with images, which are visually legible for human. As the CS measurement using SPC do not have any visually legible details, a single iteration reconstruction is employed before CNN training. The single iteration reconstruction step is required to get a minimum visually legible image that to be trained in CNN. This paper [19] also discussed on the possibility of learning the SPC mask as a future scope. Another work [3] demonstrates the process of learning the mask jointly with model weights for a 28 × 28 MNIST handwritten dataset.

In the present proposed work, it is planned to use reduced pixel transform of the training images, which can be obtained by applying a binary mask on the training, and test images and use only P% of the pixels in the image. This can be consider as CS measurement and well demonstrated in our previous work [21] using a novel technique called Genetic Algorithm-based Compressed Learning (GACL), which will increase the CNN's training accuracy by learning the binary mask. The key merits of this proposal is based on the fact that the reduced pixel resultant images retain the visual legibility even after applying the binary mask and thus the single

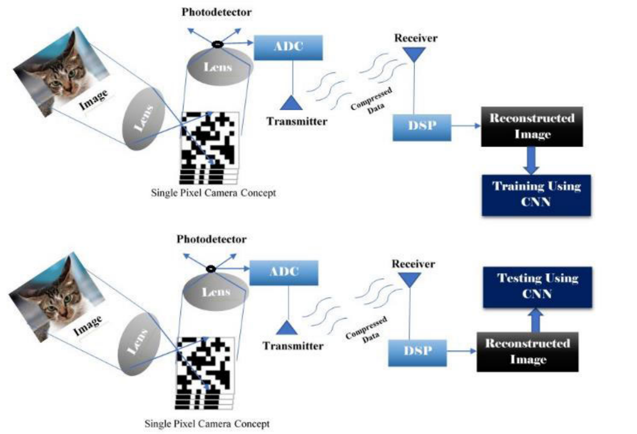


FIGURE 1. Architecture of dynamic range image classification process using SPC.

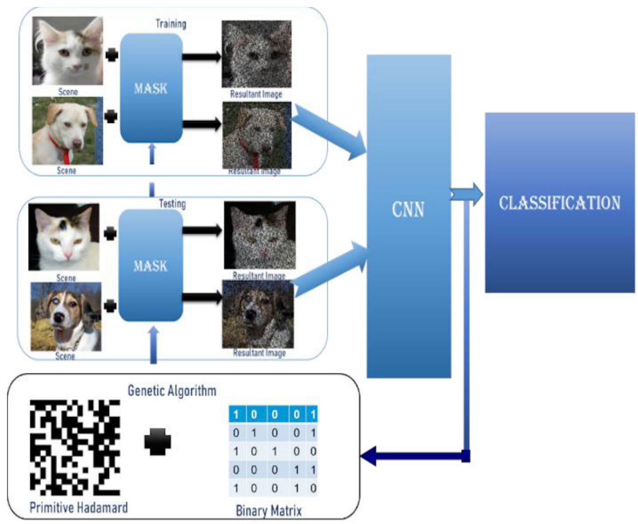


FIGURE 2. Architecture of genetic algorithm based compressive learning.

iterative reconstruction process is not required. Moreover the practical implementation of SPC is very complicated and implementation of reduce pixel transform can be realized by using simple masking procedure [21]. In this, applying binary mask is the sensing process of CS framework and this sensing process reduce the MR but the images will be still legible and can be directly used for training process as shown in Figure 2. To learn the best binary mask, GACL framework uses two different crossover process vertical crossover and diagonal crossover.

The goal of the GACL is to determine the best mask that to be applied to all of the training set images in order to increase CNN's training accuracy in compressed learning. This is accomplished by considering the CNN training accuracy to be the objective function that is to be maximized in GACL algorithm to get the best mask. It is shown that the classification accuracy for the case of retaining 10 % of the image pixels by applying binary mask is 70 % and after using

GACL this accuracy increases to 85 % which is remarkable for such a low MR [21].

In this present paper, it is proposed to extend the idea of GACL in wavelet domain. Wavelet is the best tool to compress images and masking of wavelet coefficients can be considered as a CS sensing process, which introduces further reduction of MR. In this paper, it is proposed to use the GACL learning for the mask applied to approximation wavelet coefficient and to study the learning performance of CNN and the wavelet selected for this study is the basic Haar wavelet. In the numerical demonstration, Haar wavelet is applied on the training images and approximation coefficients are used for CNN learning. In this novel approach, the reduction of MR is twofold firstly by using only the approximation coefficient of the image and secondly by masking the approximation coefficient matrix. In spite of the fact that the approximation coefficients constitutes only one fourth of the total number of pixels, the CNN learning works fine as the Haar approximation coefficient matrix visually resembles the original image and the study shows that GACL learned mask that consists of 10 % of approximation coefficients yield an accuracy of 96 %. This enables drastically low MR of 10 % of one-fourth of the total number of pixels of the original image. Thus, the methodology of GACL based compressive domain learning using wavelet coefficient will drastically reduce the need for huge data in both training and testing phases.

To reason out the increase in CNN model accuracy for the case of GACL the visual quality of the masked images are computed by calculating the measure of deviation of the masked image from the original unmasked image. The Structural Similarity Index (SSIM) and Peak signal to noise ratio (PSNR) are the good metrics to evaluate how similar two images are and in this work S_{avg} and $PSNR_{avg}$ is calculated for each member of unmasked dataset with corresponding member of masked dataset and averaged out to get average SSIM between original dataset and masked dataset. In applying mask to retain, only P% of the original image it is observed that S_{avg} decreases when P decreases and it is found interesting that S_{avg} is improved when GACL is applied. This brings out the fact that GACL learns the best mask, which will output an image perceptually very closer to the original unmasked image. The research contribution in this paper is to apply reduce pixel transformation for approximation wavelet coefficients with best-learned mask using GACL to get the best accuracy and to reason out the increase in accuracy with the help of SSIM measurements.

To summarize, the motivation of the work is to demonstrate CNN image classification by using only P% of the pixels of the original image in its wavelet domain and to achieve this following are the contributions of the paper.

- A new dataset is created by applying wavelet transform to the original dataset images and retaining only the approximation coefficient images. For the proof of concept, Haar wavelet is used and the level of decomposition is 1.
- A structured binary mask is applied to the images of the new dataset to retain only the P% of the pixels. For the proof of concept simple PWH matrix based mask is used.
- A study of performance of wavelet based compressed domain learning is done by using clean/masked training dataset and masked test dataset by observing the CNN learning accuracy. The visual quality of the images for various P values are quantified using average SSIM of the dataset and studied.
- GACL is implemented to learn the binary mask as shown in figure 2 for various P and the visual quality of the dataset images and the learning performance of CNN classification are studied for both vertical and horizontal crossover methods to show the merits of vertical crossover method.
- The experiment is demonstrated for very low MR rate of 2.5 %.

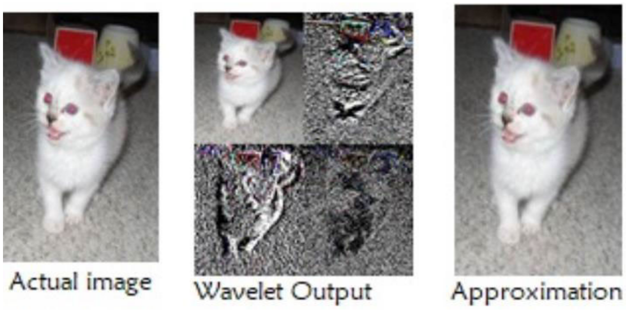
This study is structured as follows, section II discussed about compressed domain learning using Haar wavelet, section III Genetic Algorithm based Compressed Learning (GACL) using Wavelet Approximation, section IV result and discussion, lastly conclusion.

II. COMPRESSED DOMAIN LEARNING USING HAAR WAVELET

In order to compress the images Haar wavelet is applied on the dataset and the approximation coefficients are used for further process. This section divided into two, section (a) discusses about the compressed dataset and the section (b) study of performance of wavelet-based compressed domain learning.

A. COMPRESSED DATASET

As discussed in introduction the motivation of the work is to demonstrate CNN image classification by using only $\boldsymbol{P}\%$ of the pixels of the original image in its wavelet domain. This is achieved by applying wavelet transform to the dataset images to create a new dataset retaining only the approximation coefficients images. These approximation coefficient images are further masked to retain only $\boldsymbol{P}\%$ of the pixels. Using images of cats and dogs from the Kaggle dataset, Emine Cengil et al. [22] trained and tested deep learning algorithm using the caffe package, which is frequently used for deep learning. The original raw kaggle database [23] utilized to get compressed dataset for this work is made up of 2000 training and 500 test images, each of which has an RGB value of $\boldsymbol{M} \times \boldsymbol{N}$. In this technique, all the images from the dataset are resized to $\mathbf{255} \times \mathbf{255}$ standard images and Haar based discrete wavelet transforms (DWT) is applied to get approximation coefficient matrix and detailed matrices. The approximation matrix, which is down sampled and consists of 25 % of original pixel is retained and other detailed matrices are neglected as shown in Figure 3. It can be inferred that the



(a) (b) (c)

FIGURE 3. Resultant dataset image of the Wavelet transform.

(a) (b)

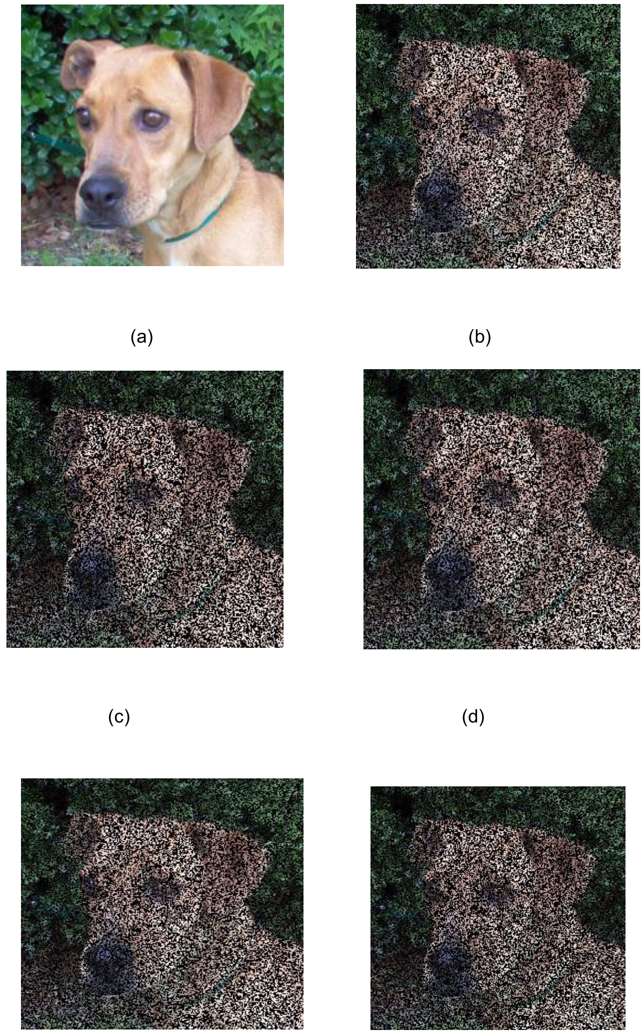
FIGURE 4. Sample masked approximation images of the dataset (a) Random matrix (b) PWH matrix.

TABLE 1. CNN model parameters [20].

Parameters	Values
Model Type	Sequential
Activation Layer	Relu
Shear Range	0.2
Zoom Range	0.2
Flip Type	Horizontal
Filter Size	64 X 64, 128X128
Kernel Size	3 X 3
Optimizer	Adam
Batch Size	32
Epoch	20-100
Class Mode	Binary
Loss function	Binary Cross Entropy
Metrics	Accuracies

approximation matrix resembles the original image and good enough to be used in CNN training.

In this proposed paper, one of the key process in obtaining the compressed domain dataset is to apply a binary mask to the images that retains $P\%$ of the pixel and makes the remaining pixels as zero. One of the direct method is to apply a random binary mask, which has $P\%$ of the pixels arbitrarily 1 and remaining 0. Using a structured binary mask



(a) (b) (c) (d) (e) (f)

FIGURE 5. Sample masked wavelet images of the dataset with P % pixel retained (a) P = 25 (b) P = 12.5 (c) P = 11.25 (d) P = 10.8 (e) P = 10.6 (f) P = 10.

like Primitive Walsh- Hadamard (PWH) mask will produce better legible image compare to the random binary mask as shown in figure 4. To get the structured binary mask, the Primitive Walsh- Hadamard (PWH) matrix, is contrasted with the binary sensing matrix. The figure 5 illustrate the masking process for various P for a sample approximation coefficient matrix.

B. STUDY OF PERFORMANCE OF WAVELET-BASED COMPRESSED DOMAIN LEARNING

The previous work [21], it is shown that the training accuracy of 97% is possible with CNN classification using the uncompressed raw original dataset [23] and the CNN architecture model parameters are provided in Table 1. In the same work, the performance of CNN is studied by reducing the pixels of the original training and testing image using a binary mask.

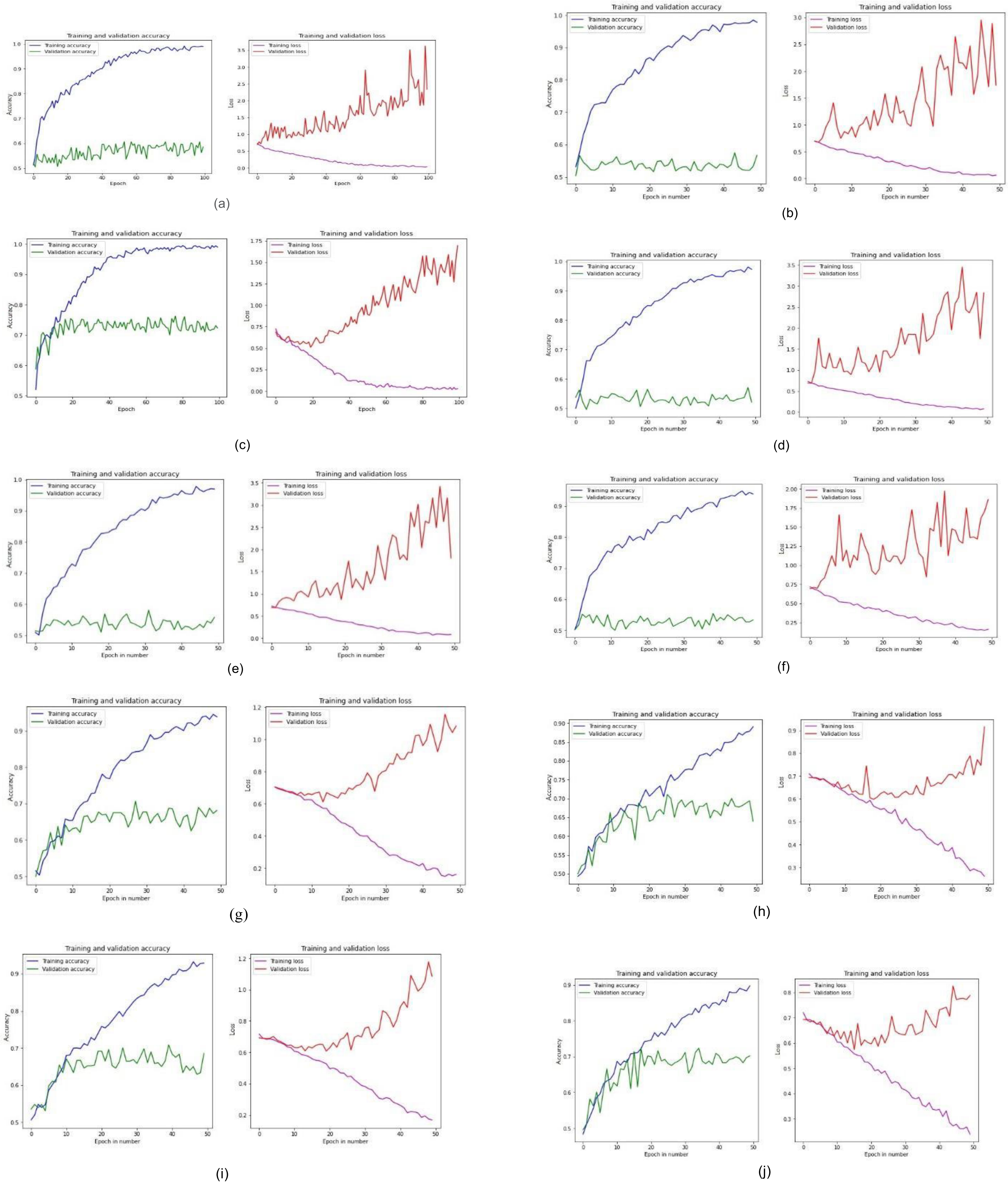


FIGURE 6. Training and validation accuracy and loss graph with respect to epochs for PWH masked dataset with P % pixels retained in testing dataset and Q % in the training dataset, (a) P = 25, Q = 25, (b) P = 25, Q = 12.5 (c) P = 25, Q = 11.25 (d) P = 25, Q = 10.8(e) P = 25, Q = 10.8 (f) P = 25 Q = 10 (g) P = Q = 11.25 (h) P = Q = 10.8 (i) P = Q = 10.6, 8 (j) P = Q = 10.

This paper proposed an experiment to conduct the similar study using the Haar wavelet approximation images. Two different ways of the learning experiment carried out for approximation coefficient images by using:

(i) Unmasked images for training and the PWH masked for testing.
(ii) PWH masked images for training and the PWH masked images for testing. The results are provided in table 2.

In these cases, the new arrived masked dataset is of low measurement rate and the image quality is measured using average SSIM and average PSNR. Let us assume $\boldsymbol{Ds}$ is the image set 1, which is original dataset, comprises of Np images and let $\hat{\boldsymbol{Ds}}$ be the image set 2, which is compressed dataset. The average SSIM ($\boldsymbol{S_{avg}}$) is calculated using the following equation (1) [24], [25].

$$\mathrm{S_{avg}} = \frac{1}{\mathrm{Np}} \sum_{\mathrm{k}=1}^{\mathrm{Np}} \mathrm{SSIM}(\mathrm{Ds}(\mathrm{k}),\hat{\mathrm{D}}\mathrm{s}(\mathrm{k})) \tag{1}$$

where, $Ds(k)$ is $\mathrm{k^{th}}$ image of the dataset.

In addition, the average PSNR ($PSNR_{avg}$) is calculated using the equation (2) [25], [26].

$$\mathrm{PSNR_{avg}} = \frac{1}{\mathrm{Np}} \sum_{\mathrm{k}=1}^{\mathrm{Np}} \mathrm{PSNR}(\mathrm{Ds}(\mathrm{k}),\hat{\mathrm{D}}\mathrm{s}(\mathrm{k})) \tag{2}$$

Case (i) here the approximation images are used for training and the PWH masked approximation images are used for testing, for $\boldsymbol{P} = \mathbf{100}$ the obtained training accuracy was 97% which is adequately good which implies no information loss in testing set. Figure 6(a) displays the model accuracies and loss function for $\boldsymbol{P} = \mathbf{100}$. As a continuation for case (i) the experiment was repeated for masked test dataset with various values of P and it was found that the training accuracy is constant and validation accuracy kept decreasing. Figure 6 (b-e) shows the model accuracies and loss function. Similarly, for case (ii) both the training and testing datasets were masked and the experimental result demonstrated that the validation accuracy was better than case (i). It should be noted that the training accuracy is 89 % and validation accuracy is 62 % when P=10 % which is the least MR in this experiment. Figure 6 (g-j) displays the loss functions and CNN model accuracies. The image quality with respect to $\boldsymbol{P}$ (% of the pixel retained) in terms of average SSIM and average PSNR using (1) and (2) is also studied and given in table 2. It can be observed from the table that when $\boldsymbol{P}$ reduce the $\boldsymbol{S_{avg}}$ and $\boldsymbol{PSNR_{avg}}$ also reduce which effects the CNN training accuracy. Figure 7 (a-b) shows the comparison graph of the accuracy with respect to percentage of pixel retained image for both case (i) and case (ii). Figure 7(c) shows the comparison graph of image quality measurements with respect to P.

III. GENETIC ALGORITHM BASED COMPRESSED LEARNING (GACL) USING WAVELET APPROXIMATION

In the experiment described in Section II, it was learnt that a CNN could recognize a wavelet approximation image with a validation accuracy of 62 % and learn from 10 % of distributed pixels in the training set with a training accuracy of 89 %. 90 % of the pixels in the mask are 0 and only 10 % of the mask are 1 and there is lot of combinations are possible for the distribution of 1s among 0s. As discussed in [21] there is

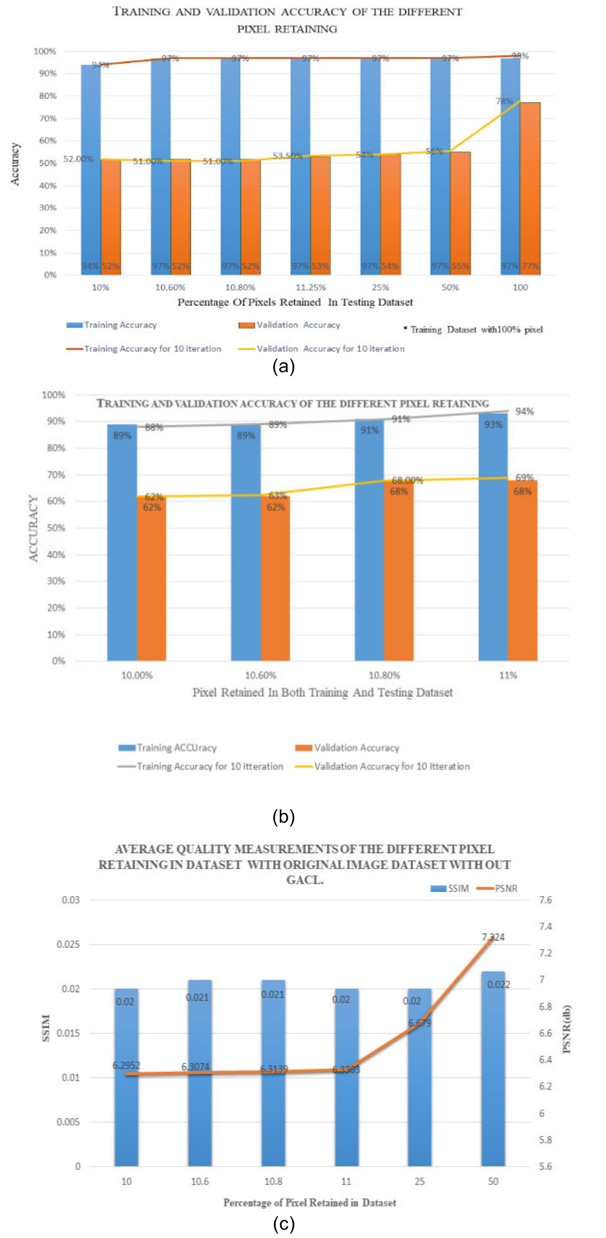


FIGURE 7. Comparison of (a) training and validation accuracy with unmasked training and masked testing dataset, (b) training and validation accuracy with masked training and masked testing dataset, (c) average quality measurement of the different pixel retaining without GACL.

lot of scope to learn the best mask using heuristic algorithm. The GACL algorithm is successfully demonstrated for MR of 10 % of original spatial domain image.

In this work, the GACL need to be tested for the approximation image whose pixel size is one fourth of the original spatial domain image. The generalized block diagram for the workflow of this approach for the training and testing the models is presented in Figure 8 for the Genetic Algorithm-based Compressed Learning (GACL) for a two-class dataset. It should be noticed that all of the images in the training and testing data sets use the same binary mask. The training for

TABLE 2. Training and validation accuracy of the different pixel retaining.

Percentage Of Pixels Retained In Approximation Images		Accuracy for single iteration		Average Accuracy For 10 Iteration		Average Quality Measurements	
Training Dataset	Testing Dataset	Training Accuracy	Validation Accuracy	Training Accuracy	Validation Accuracy	SSIM	PSNR(db)
100 %	100 %	97 %	77 %	98 %	77.82 %	1	-
100 %	50 %	97 %	55 %	97 %	55.52 %	0.022	7.3246
100 %	25 %	97 %	54 %	97 %	53.9 %	0.020	6.679
100 %	11.25 %	97 %	53 %	97 %	53.5 %	0.020	6.3303
100 %	10.8 %	97 %	52 %	97 %	51 %	0.021	6.3139
100 %	10.6 %	97 %	52 %	97 %	51 %	0.021	6.3074
100 %	10 %	94 %	52 %	94 %	52 %	0.021	6.2952
11.25 %	11.25 %	93 %	68 %	94 %	69 %	-	-
10.8 %	10.8 %	91 %	68 %	91 %	68 %	-	-
10.6 %	10.6 %	89 %	62 %	89 %	62.51 %	-	-
10 %	10 %	89 %	62 %	88 %	62 %		

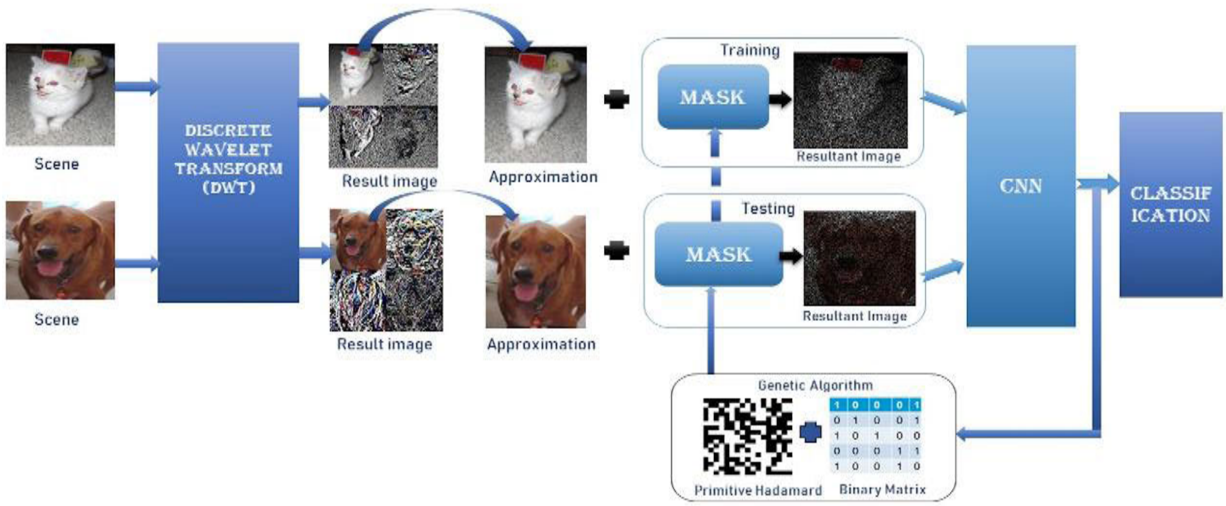


FIGURE 8. Proposed Architecture diagram of Genetic Algorithm based compressive learning using wavelet.

TABLE 3. Validation and training accuracy for vertical and diagonal crossover.

Method	Percentage of Pixels Retained		Accuracy for single iteration		Average Accuracy For Ten Iteration	
	Training Dataset	Testing Dataset	Training Accuracy	Validation Accuracy	Training Accuracy	Validation Accuracy
Vertical Crossover	10 %	10 %	93 %	58 %	92.48 %	57 %
Diagonal Crossover	10 %	10 %	96 %	58 %	96.17 %	58.2 %

this study begins with the widely used Cat and Dog datasets. One of the strengths of this work is that, in addition to training the original binary sensing matrix, applied two types of fixed binary sensing patterns.

The complete process of GACL algorithm is given in figure 9 and its pseudo code in algorithm 1. As shown in figure 9, the main objective of the algorithm is to select the best mask that to be applied to the selected dataset to

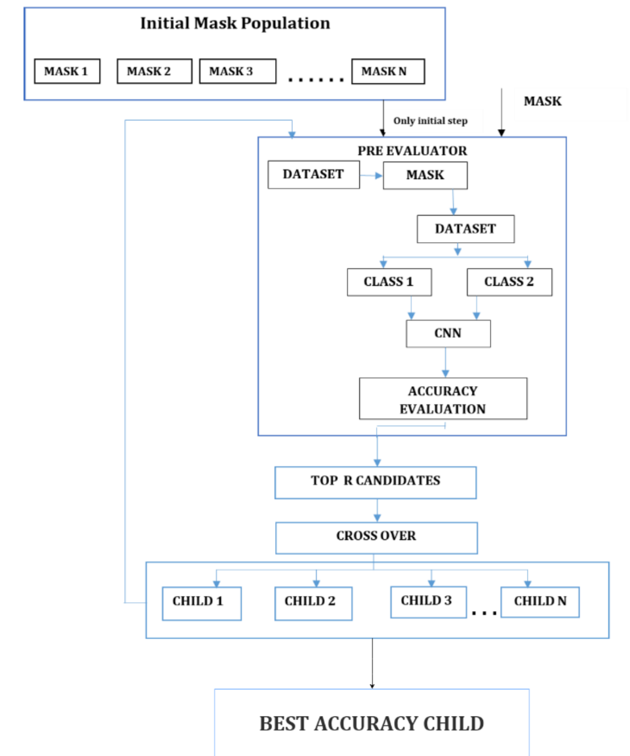


FIGURE 9. Proposed algorithm of genetic algorithm based compressive learning.

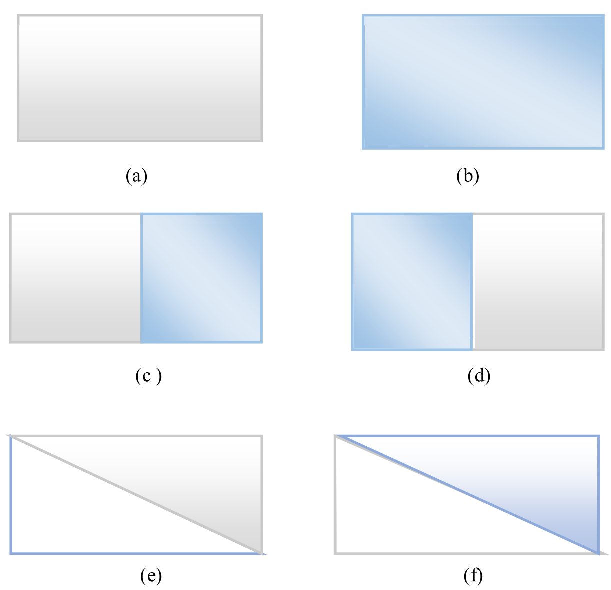


FIGURE 10. Crossover pattern of the resultant masked image (a) and (b) top two selected chromosomes (c, d) vertical crossover (e,f) diagonal crossover.

get the best training accuracy. The mask generation process is performed by two consecutive stages, a PWH masking stage followed random binary masking to get the required percentage ($\boldsymbol{P}$) of 1's, in this case $\boldsymbol{P}$ is 100. Using this process an initial mask population is created that has N masks. These masks are applied to both training and testing samples of the selected dataset and CNN learning is performed. The vertical

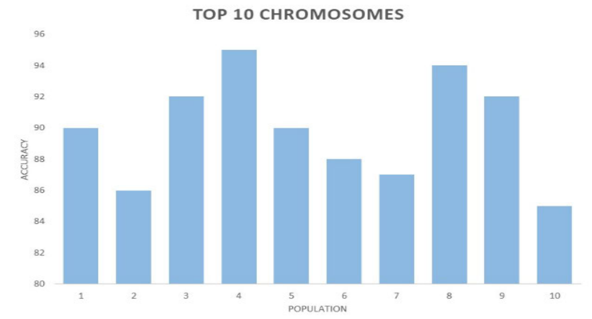


FIGURE 11. Top 10 the selected accuracies for the GACL mask.

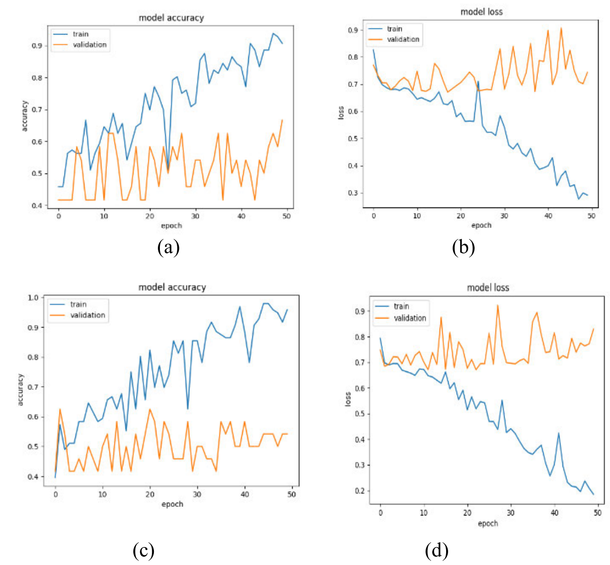


FIGURE 12. The training and Validation performance CNN model achieved through GACL for P = Q = 10 %, (a) accuracy with vertical crossover, (b) loss with vertical crossover, (c) with diagonal crossover, (d) loss with diagonal crossover.

cross over is performed by merging the left half of the first mask with right half of the second mask and vice versa and the vertical crossover pattern is shown in Figure 10 (c-d). The diagonal crossover involves merging of the diagonal region of the masks as shown in Figure 10(e-f). The repeated iterations of this process will yield the best child mask. It is worthwhile to mention here that the key bottleneck of the process that took long computation time is the CNN training that to be performed for every mask. For a proof of concept, here N and M values were considered as N = 10 and M=2. The result obtained for the first iteration is taken for the analysis. The top 10 accuracy is depicted in figure 11 and the results obtained for the best child for both types of crossover is shown in table 3. The resulting model accuracy graphs are given in Figure 12.

IV. RESULTS AND DISCUSSION

This research is all about to reduce the measurement rate of an image that to be tested for recognition of the class that the

Algorithm 1 Pseudo Code for GACL Algorithm

```
ip=size_initial_population
M = 256                  # no of rows
N=256                   # no of columns
per_1 = 0.2           #% of non-zero elements
src = "dataset_path"
for i=1:ip
   bm=rand_mask(M,N,per_1)
   h1=hadamard(M,N)
   mask=bm.*h1
   acc=pre_eval(mask,M,N,src)
   cum_acc[i]=acc
   mask_store(i,:,:)=mask        # mask of every iterations are
            stored
end
tops=sort_top_R(cum_acc,mask_store) # top R accurate masks
childset1=vertical_crossover(tops # R child by vertical crossover
childset2=diagonal_crossover(tops) #R child by diagonal crossover
for i=1:R
   mask = childset1(i,:,:)
   acc=pre_eval(mask,M,N,src)
   a1(i)=acc
   mask = childset2(i,:,:)
   acc=pre_eval(mask,M,N,src)
   a2(i)=acc
end
analyse(max(a1))
analyse(max(a2))
def pre_eval(mask,M,N,path):
   img=cv.imread(path)
   img1=cv.resize(img)
   output= img1*mask
   cv.imwrite(dest_path)
   acc=CNN(dest_path)
   return(acc)
```

image belongs. This concept being successfully demonstrated in the previous research [21] using a cat dog dataset by retaining only 10 % of the pixels of the images of both training and testing dataset and also being shown that the best mask can be learnt using genetic algorithm and the process is named as GACL. In this work, the idea is extended by using only the approximation coefficient image of the dataset samples.

The experiment is conducted for same dataset used in the above referred work [21] using haar wavelet approximation. Apparently, using $\boldsymbol{P}\%$ of the approximation implies $\boldsymbol{P}x\left(\frac{1}{4}\right)=\left(\frac{P}{4}\right)$ % of the original image and the final $\left(\frac{P}{4}\right)$ % of the pixels are assumed as compressed measurements.

Table 2 demonstrates the training and validation accuracy achieved using the CNN with parameters shown in table 1 for two different scenarios. In both the scenarios $\boldsymbol{P}\%$ of the approximation coefficient images are used for testing. The key difference is 100 % of the approximation coefficient images are used for training in scenario 1 and $\boldsymbol{P}\%$ of the approximation coefficient; images are used in scenario 2. It is important to notice from table 2 that the scenario 2, as expected training accuracy is lesser but the validation accuracy is better as the training samples has the same measurement rate like testing samples. Further, in this process a novel Genetic Algorithm-based compressive learning (GACL) method is proposed to learn the PWH mask in order to increase the model training accuracy.

The experiment is conducted for the GACL for the case of measurement rate (MR) of 10 % by keeping just 10 % of the pixels in every image in both the training and testing dataset that represents two classes and the results are tabulated in table 3. From the table 3 it can be observed that the training accuracy is increased from 89 % to 93 % by employing vertical crossover in the offspring formation of GACL From the same table, it can also be observed that when the diagonal crossover is used in place of vertical crossover the training accuracy increased to 96 % from 93 %. As shown in table 4, this training accuracy of 96 % achieved in wavelet domain with an apparent measurement rate of 10/4 is much better than the accuracy achieved in prior work [21] in which training accuracy is 85 % using a spatial domain dataset with the measurement rate of 10. It should be noted training accuracy of 96 % is achieved with the very less MR of 2.5 % when using wavelet approximation. It is observed that the average SSIM and average PSNR reduces as $\boldsymbol{P}$ reduces, and it is interesting to note that average SSIM and average PSNR improves when GACL is used as shown in table 5. It is observed that the average SSIM for the original image set 1 with 100 % and masked image set 2 with 50 % is 0.022 and it is improved to 0.1735 for vertical crossover and 0.2890 for diagonal crossover.

TABLE 4. Comparison table on spatial domain and frequency domain.

Percentage of Pixels Retained		Without wavelet		With wavelet	
Training Dataset	Testing Dataset	Training Accuracy	Validation Accuracy	Training Accuracy	Validation Accuracy
100 %	100 %	97 %	76 %	97 %	77 %
10 % (without GACL)	10 %	77 %	61 %	89 %	62 %
10 % (with GACL)	10 %	85 %	62 %	96 %	62 %

Similarly, average PSNR also improved from 7.324 to 10.5676 for diagonal crossover. Average SSIM and Average PSNR are calculated for 25 % and 10 % of the masked dataset with the original unmasked dataset is tabulated in table 5,

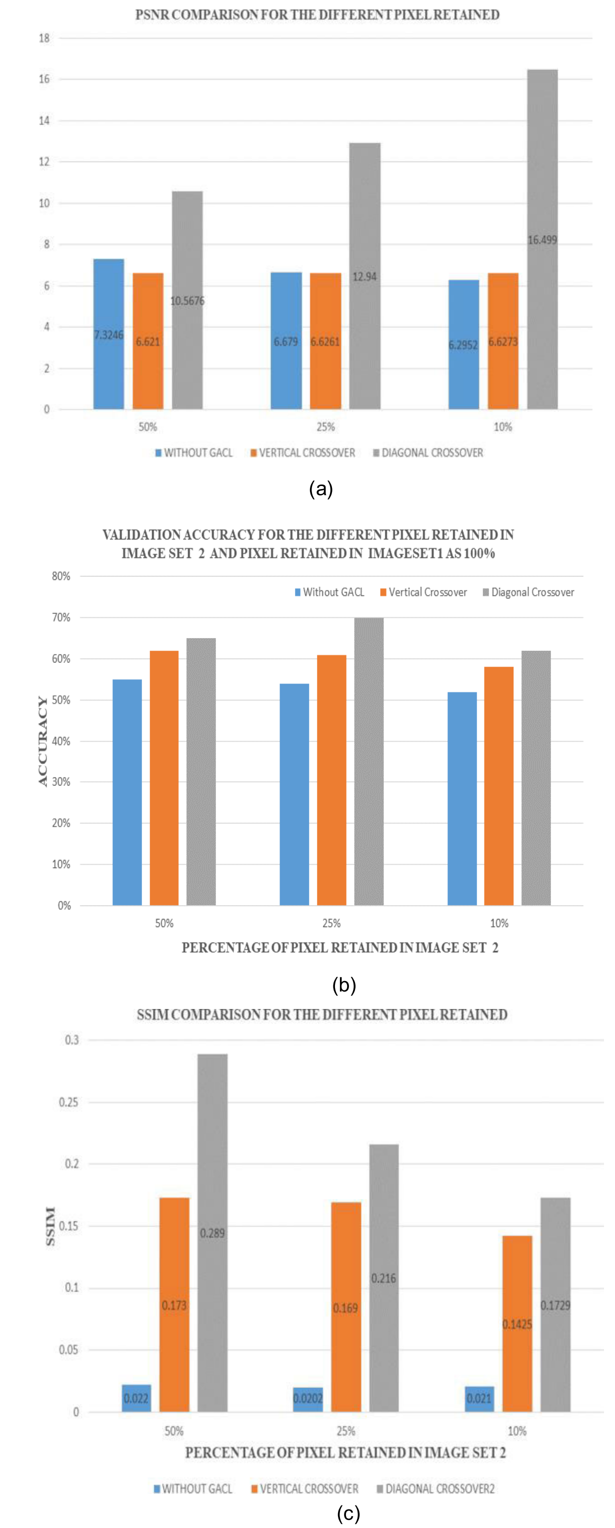


FIGURE 13. **Comparison of (a) training and validation accuracy with unmasked training and masked testing dataset, (b) average PSNR measurement, (c) average SSIM measurement of the different pixel retaining with GACL.**

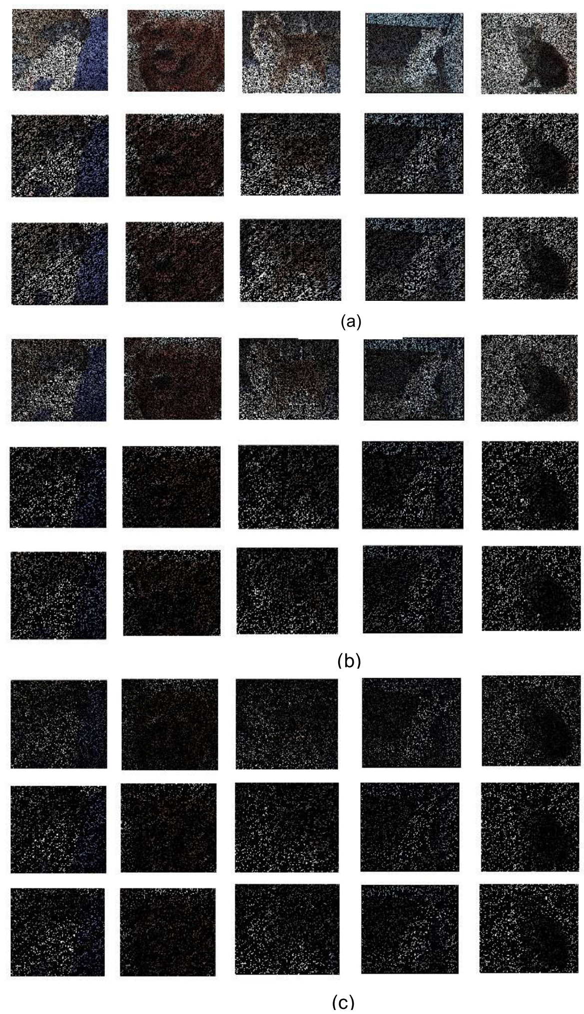


FIGURE 14. **Intermediate resultant images for 5 selected samples images (a) 50 % of pixel retained images row 1: without GACL, row 2: vertical crossover, row 3: diagonal crossover, (b) 25 % of pixel retained images row 1: without GACL, row 2: vertical crossover, row 3: Diagonal crossover, (c) 10 % of pixel retained images row 1: without GACL, row 2: vertical crossover, row 3: Diagonal crossover.**

and the comparison graphs are as shown in figure 13. The intermediate results for 5 sample images for various $\boldsymbol{P}$ values used in table 5 for the 3 different cases without GACL, GACL with vertical crossover, GACL with diagonal crossover is depicted in figure 14 for subjective analysis. From figure 14 it can be observed that the contribution of GACL to make the masked images more legible is more significant when $\boldsymbol{P}$ is less i.e. very low measurement rate.

The time taken for executing the entire experiment using python programming on 11th Gen Intel Core i7 processor with 2.80 GHz clock frequency and 16.0 GB internal RAM is shown as table 6. The computational complexity in GACL is mainly because a long CNN training procedure is involved in every iteration to determine the performance matrix of the mask.

As a summary, following points can be inferred from the results of the numerical experiment.

(1) Training wavelet approximation images on CNN are as good as using original images. ($\boldsymbol{P} = \mathbf{100}$, Training accuracy 97 % and validation accuracy 77 %, table 2)

TABLE 5. Quality comparison table for the different pixel retained.

Method	Accuracy			Average Quality Measurements	
	Image set	Training	Validation	SSIM	PSNR(db)
Without GACL	50 %	97 %	55 %	0.022	7.3246
Vertical Crossover	50 %	96 %	62 %	0.1735	6.6210
Diagonal Crossover	50 %	95 %	65 %	0.2890	10.5676
Without GACL	25 %	97 %	54 %	0.020	6.679
Vertical Crossover	25 %	96 %	61 %	0.1649	6.6261
Diagonal Crossover	25 %	94 %	70 %	0.2162	12.9457
Without GACL	10 %	94 %	52 %	0.021	6.2952
Vertical Crossover	10 %	94 %	58 %	0.1425	6.6273
Diagonal Crossover	10 %	96 %	68 %	0.1794	16.499

TABLE 6. Time measurement table for the different experiments (I7P, 2.80GHZ frq, 16GB ram).

Experiment	Time Duration in minutes
Image Masking process	4 minutes
CNN accuracy measurement Without GACL(single iteration)	10.2 minutes
CNN accuracy measurement With GACL(single iteration)	17. minutes

(2) The $\boldsymbol{S_{avg}}$ reduces when the test wavelet images are masked and validation accuracy reduce. ($\boldsymbol{P} = \mathbf{10}$, validation accuracy 58 %, $\boldsymbol{S_{avg}} = \mathbf{0.022}$, table 2)

(3) When training is done on masked images, validation accuracy for masked test wavelet images improves. ($\boldsymbol{P} = \mathbf{10}$, validation accuracy 62 %, $\boldsymbol{S_{avg}} = \mathbf{0.022}$, table 2).

(4) The DCNN training on masked images are usually characterized by lower training accuracy. ($\boldsymbol{P} = \mathbf{10}$, training accuracy 89 %, $\mathbf{S_{avg}} = \mathbf{0.22}$, table 2

(5) GACL learning of binary mask improves SSIM which improves the learning accuracy. ($\boldsymbol{P} = \mathbf{10}$, vertical cross over GACL, training accuracy 93 %, $\boldsymbol{S_{avg}} = \mathbf{0.1425}$, table 5)

(6) The best SSIM and learning accuracy is achieved when diagonal crossover is used along with GACL. ($\boldsymbol{P} = \mathbf{10}$, Diagonal cross over GACL, training accuracy 96 %, $\boldsymbol{S_{avg}} = \mathbf{0.1794}$, table 5)

V. CONCLUSION

In this paper, a performance study of DCNN based two classes image classification problem is performed using only the approximation Haar wavelet coefficient images of the standard cat and dog data set images relying on the fact that an approximation image retains the visual information of the original image. The study is further performed by applying a CS sensing matrix, which is a PWH, based binary mask on the approximation images to retain only P% of the pixels of the approximation images and in order to measure the image quality average SSIM (S_{avg}) and average PSNR ($\mathrm{PSNR_{avg}}$) are calculated for the new derived dataset. The approximation image has only $1/4^{th}$ of the original image and retaining only P% of pixels implies that only (P/4) % of the pixels of the original image are used apparently. When the DCNN training is done on clean un masked approximation images of a cat-dog dataset the classification training accuracy is found to be 97 % and when the CNN model is tested with masked test sets for the case of 10 % of pixels retained, the S_{avg} falls to 0.022, the validation accuracy falls up to 58%. When the DCNN training is done on masked cat-dog dataset, as expected the classification training accuracy of the model falls up to 83 % but the validation accuracy increases up to 62 % for the 10 % retained test dataset. The increased validation accuracy is due to the fact that the training set is masked alike testing set and generalize the test set better. This is referred as compressive learning in which learning is directly done on compressed measurements, which will minimize the measurement rate of the sample images both for training and as testing. The performance of both the scenarios are studied by varying the percentage of pixels retained ($P\%$) and reported. The mask used for all the images in both the training and testing dataset are same. As a novel idea, it is proposed to learn the best mask to be applied to the approximation coefficient image by learning the points where the mask should be 1 and where it to be 0 by using Genetic Algorithm-based compressive learning (GACL) framework. GACL algorithm is an iterative algorithm, which uses the DCNN accuracy as the performance metrics to get the best child mask. The implementation details of GACL using vertical and diagonal crossover is discussed thoroughly with a block diagram with supporting pseudocode.

The GACL experiment on approximation images of cat-dog dataset for the case of 10 % measurement rate/compression ratio (CR) 90 % by maintaining just 10 % of the pixels in every image both in training and testing dataset increases the S_{avg} from 0.022 to 0.1425 which increases the training accuracy to 89 % for the case of vertical cross over. When diagonal crossover is applied in place of vertical crossover in GACL the S_{avg} reaches to the highest value of 0.1794 and the training accuracy increases to its maximum of 96 %. This is very good improvement and better than the training accuracy of 85 % obtained for a prior GACL work done [21] using spatial domain images. The training done on masked approximation images apparently results in a high outcome with a classification accuracy of 96 % using $\frac{P}{4} = 2.5\%$ of the total pixels of spatial domain image using the novel GACL framework with diagonal crossover, This novel GACL method has a tremendous application in an IoT framework where the measurement rate of both training and testing rate can be much minimized for very low values.

The suggested GA algorithm worked well with the existing GA-based CNN. This proposed work is just a proof of concept to show that reduce pixel transform in wavelet domain can be used in compressive domain learning for very MR. the work is limited to very basic wavelet technique and a basic structural masking technique. In the future, this work can be extended with different wavelet methods, different masking methods and more computationally intensive advanced heuristic techniques to learn the masks.

REFERENCES

[1] S. Lohit, K. Kulkarni, P. Turaga, J. Wang, and A. C. Sankaranarayanan, "Reconstruction-free inference on compressive measurements," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2015, pp. 16–24.

[2] S. Lohit, K. Kulkarni, and P. Turaga, "Direct inference on compressive measurements using convolutional neural networks," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 1913–1917.

[3] A. Adler, M. Elad, and M. Zibulevsky, "Compressed learning: A deep neural network approach," 2016, *arXiv:1610.09615*.

[4] K. Kulkarni and P. Turaga, "Reconstruction-free action inference from compressive imagers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 4, pp. 772–784, Apr. 2016.

[5] M. A. Davenport, M. F. Duarte, M. B. Wakin, J. N. Laska, D. Takhar, K. F. Kelly, and R. G. Baraniuk, "The smashed filter for compressive classification and target recognition," *Proc. SPIE*, vol. 6498, pp. 142–153, Jan. 2007.

[6] Y. Li, C. Hegde, A. C. Sankaranarayanan, R. Baraniuk, and K. F. Kelly, "Compressive image acquisition and classification via secant projections," *J. Opt.*, vol. 17, no. 6, Jun. 2015, Art. no. 065701.

[7] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[8] A. Khmag, "A robust watermarking technique for high-efficiency video coding (HEVC) based on blind extraction scheme," *Social Netw. Comput. Sci.*, vol. 2, no. 4, p. 329, Jul. 2021.

[9] R. Calderbank, S. Jafarpour, and R. Schapire, "Compressed learning: Universal sparse dimensionality reduction and learning in the measurement domain," *Preprint*, 2009.

[10] A. Khmag, "Additive Gaussian noise removal based on generative adversarial network model and semi-soft thresholding approach," *Multimedia Tools Appl.*, vol. 82, pp. 1–21, Aug. 2022.

[11] A. Khmag and N. Kamarudin, "Natural image deblurring using recursive deep convolutional neural network (R-DbCNN) and second-generation wavelets," in *Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA)*, Sep. 2019, pp. 285–290.

[12] J. Ballé, V. Laparra, and E. P. Simoncelli, "End-to-end optimized image compression," 2016, *arXiv:1611.01704*.

[13] Y. Choi, M. El-Khamy, and J. Lee, "Variable rate deep image compression with a conditional autoencoder," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 3146–3154.

[14] N. Johnston, D. Vincent, D. Minnen, M. Covell, S. Singh, T. Chinen, S. J. Hwang, J. Shor, and G. Toderici, "Improved lossy image compression with priming and spatially adaptive bit rates for recurrent networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4385–4393.

[15] E. Agustsson, M. Tschannen, F. Mentzer, R. Timofte, and L. Van Gool, "Generative adversarial networks for extreme learned image compression," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 221–231.

[16] C. Y. Wang, H. Y. M. Liao, Y. H. Wu, P. Y. Chen, J. W. Hsieh, and I. H. Yeh, "Learning in the frequency domain," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Seattle, WA, USA, Jun. 2020, pp. 1740–1749.

[17] S. A. Nossier, J. Wall, M. Moniri, C. Glackin, and N. Cannings, "A comparative study of time and frequency domain approaches to deep learning based speech enhancement," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–8.

[18] T. Guo, T. Zhang, E. Lim, M. Lopez-Benitez, F. Ma, and L. Yu, "A review of wavelet analysis and its applications: Challenges and opportunities," *IEEE Access*, vol. 10, pp. 58869–58903, 2022.

[19] Y. Xu, W. Liu, and K. F. Kelly, "Compressed domain image classification using a dynamic-rate neural network," *IEEE Access*, vol. 8, pp. 217711–217722, 2020.

[20] R. Torfason, F. Mentzer, E. Agustsson, M. Tschannen, R. Timofte, and L. Van Gool, "Towards image understanding from deep compression without decoding," 2018, *arXiv:1803.06131*.

[21] B. Ali B H and P. Ramachandran, "Compressive domain deep CNN for image classification and performance improvement using genetic algorithm-based sensing mask learning," *Appl. Sci.*, vol. 12, no. 14, p. 6881, Jul. 2022.

[22] E. Cengil, A. Cinar, and Z. Guler, "A GPU-based convolutional neural network approach for image classification," in *Proc. Int. Artif. Intell. Data Process. Symp. (IDAP)*, Sep. 2017, pp. 1–6.

[23] Accessed: Feb. 15, 2022. [Online]. Available: https://www.kaggle.com/datasets/tongpython/cat-and-dog

[24] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[25] A. Devnani and C. D. Rawat, "Comparative analysis of image quality measures," in *Proc. Int. Conf. Global Trends Signal Process., Inf. Comput. Commun. (ICGTSPICC)*, Dec. 2016, pp. 353–357.

[26] Z. Wang and C. Alan Bovik, "Mean squared error: Love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98–117, Jan. 2009.

BABA FAKRUDDIN ALI B. H. received the bachelor's degree in electronics and communication engineering from Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India, and the M.Tech. degree in computer science and engineering from the Canara Engineering College, VTU. He is currently pursuing the Ph.D. degree (Research) with the School of Electronics and Communication (SENSE), Vellore Institute of Technology (VIT).

His research interests include image processing, machine learning, artificial intelligence, data science, embedded systems, robotics, and the Internet of Things (IoT).

PRAKASH RAMACHANDRAN received the bachelor's degree in engineering from the Government College of Technology (GCT), Coimbatore, India, the master's degree in engineering (embedded systems) from the College of Engineering Guindy, Anna University, and the Ph.D. degree in compressive sensing from the Vellore Institute of Technology (VIT), Chennai.

He is currently an Associate Professor with the School of Electronics Engineering, VIT, Vellore. He is also a Tech Lead with Visteon in a speech recognition project for two years. His research interests include image processing, machine learning, artificial intelligence, data science, embedded systems, and the Internet of Things (IoT).