❒    538

# Hadamard Walsh space based hybrid technique for image data augmentation

**Vaishali Suryawanshi, Tanuja K. Sarode**
Department of Computer Engineering, Thadomal Shahani Engineering College, Bandra West, India

| Article Info | ABSTRACT |
|---|---|
| | Image data augmentation (IDA) is common when deep learning is used for image classification to address the issue of overfitting. Overfitting occurs when the datasets are small and the deep learning models have a huge capacity. Overfitting models have low training errors but high validation errors and result in poor generalization. Several methods have been researched in this context, but frequency domain-based methods are less explored. In this research, we have explored the Hadamard and Walsh space and developed two hybrid technique for IDA. The proposed techniques use a combination of Hadamard/Walsh transform and geometrical transformations. Empirical study is carried out using the VGG-16 model for image classification on the CIFAR-10 dataset and the results are compared with existing methods. The analysis of the results shows that the proposed techniques improve the evaluation parameters significantly. Further, analysis of training loss vs. validation loss shows that the proposed Hadamard-based hybrid methods have better generalization ability than the proposed Walsh-based hybrid method. |

*Corresponding Author:*

Vaishali Suryawanshi
Department of Computer Engineering, Thadomal Shahani Engineering College
TPS-III, P.G. Kher Marg, Off-Linking Road, Bandra West, Mumbai, Maharashtra, India
Email: vaishali.suryawanshi@thadomal.org

## 1.    INTRODUCTION

Deep learning research community [1]–[24] has highlighted the need of huge datasets for training high-capacity models and experimented with various image data augmentation (IDA) to address the issue of data scarcity. IDA techniques perform some transformation on the original distribution to generate synthetic images which are used to enlarge the dataset. IDA techniques are state of the art methods in applications such as COVID-19 detection [5]–[12], skin lesion detection [13]–[19], face editing [20], medical image synthesis [21], and plant disease detection [22] to improve the performance of the models.

IDA methods based on geometric transformations (GT), colour manipulation, and deep learning are common in the literature [2]–[4]. GT methods include simple operations such as translation width shift, height shift, scaling, rotation, shear, and cropping erasing. Colour manipulation techniques include adjusting the contrast, generating new images applying new color map, and adjusting brightness [3]. Deep learning methods such as generative adversarial network (GAN), variational autoencoders (VAE), StyleGAN are experimented for IDA by researchers [3], [10], [12], [20], [21]. Meethongjan *et al.* [23] proposed feature selection and colour features based data augmentation technique and noted improvement in performance. This method applied local binary patterns (LBP) to extract local texture features and fisher score to select the most useful features. Frequency domain-based methods are applied by some researchers [24], where the frequency coefficients are manipulated to generate new images.

Some of the issues with these methods are they can generate a synthetic dataset that has poor feature set and diversity, insufficient quantity, huge time required to generate the dataset, complexity of methods and their applicability may be specific. In the absence of rich feature set and diversity in the original data distribution, the data augmentation methods are most likely to generate synthetic data with poor feature set/diversity. Most of the methods can generate limited set of data, for example, GTs [3] such as horizontal flip can generate just one additional image. Such methods have to be applied in combination to generate a decently large set. Methods based on complex networks such as GAN, VAE, StyleGAN are time consuming although they can generate huge amount of realistic image data with sufficient diversity. Applicability of some methods such as GT based methods may be limited. GT such as rotation/shear operation for example can generate new images, that might change the meaning or may be irrelevant in medical image classification. For example, rotation applied on image of digit 6 might result in image 9 in digit recognition application. In medical image related applications scaling operation performed on the tumour may not result in correct diagnosis. Colour manipulation techniques such as applying new colour map on skin lesion images may lead to incorrect diagnosis.

IDA not only addresses the issue of data scarcity but also addresses the issue of overfitting which generally arises due to complex deep learning models applied on smaller datasets. Though, in the literature the problem of overfitting is addressed by applying external regularization methods [25], these methods effectively reduce the capacity of the model. IDA addresses the issue from the fundamental problem of data scarcity. While applying the IDA technique it is necessary to weigh the effectiveness of the IDA method on the generalization ability of the network.

The contribution of the research proposed in this article is two-fold. First, we have developed hybrid method using combination of frequency domain coefficients and GTs. Two methods, one that uses Hadamard coefficients and the other that uses Walsh coefficients with a combination of GTs are developed for generating synthetic images. Secondly, we have studied the effect of the proposed technique on the generalization ability of the VGG16 network. The proposed hybrid method has shown significant improvement in the performance of VGG-16 model compared to the existing methods [24] for CIFAR-10 dataset [26]. Further analysis of training loss (TL) vs validation loss (VL), we observed that the proposed methods have shown comparable improvement in overfitting whereas some IDA methods are not very effective in addressing overfitting.

## 2. THEORETICAL BACKGROUND

In this section, we have discussed the theoretical background of Hadamard transform (HT), Walsh transform and method for obtaining the latent coefficient matrix (LCM) using these transforms. We have also discussed reconstruction of an image using LCM.

### 2.1. Hardmard transform

HT is a real, symmetric, orthogonal transform that decomposes a signal into a set of orthogonal, rectangular waveforms called Walsh functions. The HT contains only +1 or -1 values and is fast and requires only $N\log_2 N$ additions/subtractions [27]. Basic Hadamard matrix H is of order 2 and is given in (1):

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{1}$$

A Hadamard matrix is always order of N where N is $2^n$ n>=2. Any higher order matrix of order $2^n$ can be generated from Hadamard of order $2^{n-1}$ as given in (2).

$$H_{2^n} = \begin{bmatrix} H_{2^{n-1}} & H_{2^{n-1}} \\ H_{2^{n-1}} & -H_{2^{n-1}} \end{bmatrix} \tag{2}$$

For example, using in (2), a Hadamard of order 4 (n=2) can be generated using Kronecker product.

$$H_{2^2} = H_4 = \begin{bmatrix} H_{2^{2-1}} & H_{2^{2-1}} \\ H_{2^{2-1}} & -H_{2^{2-1}} \end{bmatrix} = \begin{bmatrix} H_2 & H_2 \\ H_2 & -H_2 \end{bmatrix}; \ H_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \tag{3}$$

### 2.2. Walsh transform

A Walsh transform matrix of order N can be can be generated from the Hadamard matrix of order N (HN) by rearranging the rows of Hadamard as per the increasing order of sequency number of sign changes in

a row. Thus, Walsh of order 4 (W4) can be obtained from H4 as shown below. Values written outside the W4 matrix show the number of sign changes in each row.

$$W_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \end{matrix}$$

### 2.3. Obtaining latent coefficient matrix

Consider a colour image IN of size NXN. The LCM of order N for image IN can be obtained using in (4).

$$LCM_N = T_N * I_N * T_N \tag{4}$$

Here, $T_N$ can be the Hadamard matrix $H_N$ or the Walsh matrix $W_N$. We can reconstruct the original image using in (5).

$$RI_N = \frac{1}{N}[T_N * LCM_N * T_N] \tag{5}$$

Hadamard and Walsh have very good energy compaction and can be applied for IDA, filtering and design of codes [27]. The high energy coefficients are mostly concentrated in the top left corner of the LCM. Original image can be reconstructed by utilizing all the coefficients from the LCM. For generating synthetic images, we can utilize only certain percentage of the latent coefficients. The quality of the reconstructed image depends on the percentage of the high energy coefficients we have utilized during the reconstruction process.

## 3.    PROPOSED HYBRID TECHNIQUE AND IMAGE CLASSIFICATION METHODOLOGY

In this section, a hybrid method for generating synthetic images is discussed. We propose hybrid technique that uses Hadamard/Walsh transform and geometrical transforms for IDA. The methodology for image classification using popular convolutional neural network architecture VGG16 model is also briefly discussed here.

### 3.1. Proposed hybrid technique

The proposed hybrid technique follows 3 stages viz. generating synthetic images, applying multiple geometric transformations (MGT) on the images, data augmentation process for enlarging the dataset. The hybrid technique developed using MGT and Hadamard is MGT_original base with Hadamard basis (OBHB). The hybrid technique developed using MGT and Walsh is MGT_original base with Walsh basis (OBWB).

### 3.1.1. Generating synthetic images

First the colour image are separated in different colour channels. Each colour channel is transformed using the selected transform (Hadamard/Walsh) using in. (4) to obtain the LCM for each colour channel. Utilizing only 70% of the high energy coefficients of these LCMs, we have reconstructed each channel using in (5). These reconstructed channels are used to form the synthetic image. This synthetic image retains the most relevant features from the original image while modifying the low frequency components. Selecting very few high energy coefficients may result in loss of relevant features giving poor quality image and might not be suitable for tasks such as skin lesion detection.

### 3.1.2. Applying multiple geometric transformations

Here, MGT are applied on the image to generate new synthetic images. In this experiment, we have applied following transformations on each image and generated 4 new images: rotation=20, width shift=0.1, height shift=0.1, and horizontal flip=true.

### 3.1.3. IDA process for enlarging the dataset

In the data augmentation process, we enlarge the dataset by adding following sets of images.
a)    Original batch images (OB): Images in the original training set.
b)    Reconstructed images (RI): Synthetic images generated after applying Hadamard/Walsh transform as explained in section 2.3.
c)    MGT_OB: Synthetic images generated after applying MGT on OB images.
d)    MGT_OBRI: Synthetic images generated after applying MGT on the images created in RI (set 2). Replace RI by WB while referring to Walsh Batch and by HB while referring to Hadamard Batch images.

IDA using the proposed hybrid technique for Walsh transform is explained in Figure 1. Here, Walsh image (WI) is the synthetic image generated after applying Walsh transform on original image I. Using the proposed technique, single image I can generate 9 synthetic images using proposed hybrid technique as explained:

{Original image I, Walsh Image (WI), I_rotate, I_width_shift, I_height_shift, I_flip, WI_rotate, WI_width_shift, WI_height_shift, WI_flip}
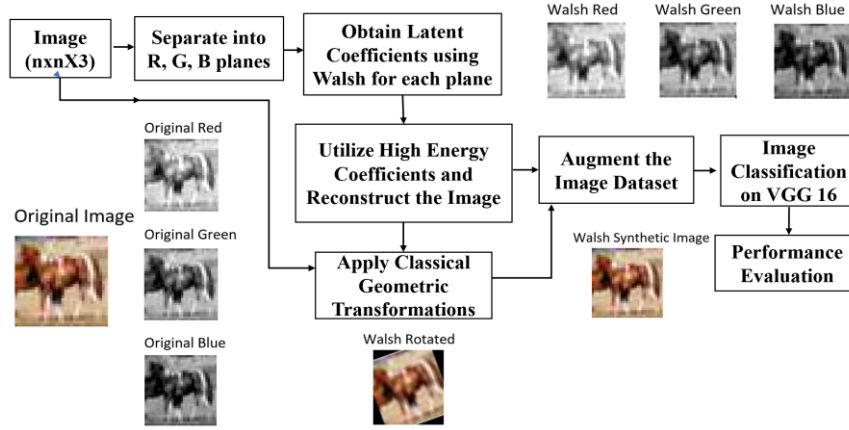


Figure 1. Proposed hybrid technique for IDA using walsh LCM

## 3.2. Image classification methodology

The image classification approach using deep learning-based model is briefly described here. First, we divide the original dataset into training and testing sets in the ratio of 80:20 and applied the proposed IDA technique on the training set to enlarge the dataset. The VGG16 [28] model is then trained with the augmented dataset followed by hyperparameter tuning. Note that the test set is not augmented. Performance evaluation for multiclass classification problem with N classes is done using the evaluation parameters computed using confusion matrix as given by in (6) to (12). The usual meaning and interpretations of the evaluation parameters are common [29], [30] and not described here.

a) Accuracy (Ci):

$$Accuracy(C_I) = \frac{\sum_{i=1}^{N} TP(C_i)}{\sum_{I=1}^{N} \sum_{j=1}^{N} C_{i,j}} \quad (6)$$

b) True positive rate (TPR)/Recall:

$$TPR(C_i) = \frac{TP(C_i)}{TP(C_i) + FN(C_i)} \quad (7)$$

c) False positive rate (FPR):

$$FPR(C_i) = \frac{FP(C_i)}{FP(C_i) + TN(C_i)} \quad (8)$$

d) Weighted recall (WR):

$$WR = \frac{\sum_{i=1}^{N} TP(C_i)}{\sum_{i=1}^{N} [TP(C_i) + FN(C_i)]} \quad (9)$$

e) Weighted precision (WP):

$$WP = \frac{\sum_{i=1}^{N} TP(C_i)}{\sum_{i=1}^{N} [TP(C_i) + FP(C_i)]} \quad (10)$$

f) Weighted F1-score:

$$WF1 = 2 * \frac{WR*WP}{WR+Wp} \qquad (11)$$

g)  Matthew's correlation coefficient (MCC) [-1, +1]:

$$MCC = \frac{c*s - \sum_k^K p_k * t_k}{\sqrt{(s^2 - \sum_k^K p_k^2)*(s^2 - \sum_k^K t_k^2)}} \qquad (12)$$

Where c: Total number of samples correctly predicted, s: Total number of samples, tk: number of times class k truly occurred, pk: number of times class k was predicted, and Ci: ith Class. MCC values are normalized so that the range is [0,1].

h)  Area under curve (AUC):

The AUC curve is a plotted a graph of TPR vs FPR for a multiclass classification problem. Here, N number of AUC-ROC graphs are plotted (one for each class), where the chosen class is positive and remaining others form a negative class

## 4.  RESULTS AND DISCUSSION

In this section, the description of original CIFAR-10 dataset, augmented dataset using the proposed method and the results are presented. The experimental set-up used in this research is described here. Investigations are carried out on the CIFAR-10 dataset using VGG16 architecture for image classification. The Google Colaboratory (Google Colab) using a GPU processor environment was utilized. The Octave environment was used for manipulating the Hadamard /Walsh space of the images from the dataset. The model is trained with SGD optimizer with a learning rate of 0.001 with categorical cross entropy as loss function.

### 4.1. CIFAR-10 dataset

The CIFAR-10 dataset [26] consists of a total of 60000 32x32 colour images in 10 clas-ses (C1: airplane, C2: automobile, C3: bird, C4:cat, C5: deer, C6: dog, C7: frog, C8: horse, C9: ship, C10: truck), with 6000 images per class. There are 50000 training images and 10000 test images. The 50000 training images are divided into 5 batches each with 10000 randomly chosen images.

### 4.2. Creating augmented dataset using the proposed technique

The proposed techniques MGT_OBHB and MGT_OBWB (as discussed in section 3.1) are applied on each of the CIFAR-10 batches. Using each proposed technique an enormous dataset of 100000 images is created from each batch of 10000 images (Total 500000 images using 50000 training images) as explained in section 3.1.

### 4.3. Result analysis

The results of the proposed MGT_OBWB and MGT_OBHB techniques are presented in Tables 1 and 2 respectively. MGT_OBWB has performed better than the MGT_OBHB technique. The average results of the proposed techniques are compared with the IDA techniques presented in [24]. The comparison of averages of all the techniques are presented in Table 2. The proposed MGT_OBWB method has shown significant improvement in all the evaluation parameters are presented in Table 3. Note that the AUC scores and N_MCC values are almost close to each other indicating that analysis can be done by either of the two.

### 4.4. Addressing generalization

The generalization ability of the proposed methods and methods in [24] is compared using TL and VL (Refer Table 4). The analysis of TL vs VL shows that MGT_OBHB has slightly better generalization capability. Further analysis of TL and VL for OB, OBWB, and OBHB show that the corresponding models are overfitted for the given dataset. MGT_OB has improved the overfitting. Also, VL for OBWB is the highest whereas for MGT_OB it is the lowest.

Table 1. Results of the proposed MGT_OBWB technique (dataset size:100000)

| Batch | Accuracy | F1 score | AUC | N_MCC |
|---|---|---|---|---|
| B1 | 0.7272 | 0.7214 | 0.8486 | 0.8491 |
| B2 | 0.7110 | 0.7060 | 0.8392 | 0.8400 |
| B3 | 0.7284 | 0.7222 | 0.8492 | 0.8496 |
| B4 | 0.7134 | 0.7101 | 0.8407 | 0.8412 |
| B5 | 0.7400 | 0.7389 | 0.8553 | 0.8563 |
| Avg. | 0.7240 | 0.7197 | 0.8466 | 0.8472 |

Table 2. Results of the proposed MGT_OBHB technique (dataset size:100000)

| Batch | Accuracy | F1 score | AUC | N_MCC |
|-------|----------|----------|---------|--------|
| B1 | 0.7128 | 0.7105 | 0.84013 | 0.8407 |
| B2 | 0.6606 | 0.6526 | 0.8114 | 0.8125 |
| B3 | 0.6954 | 0.6952 | 0.8307 | 0.8311 |
| B4 | 0.6878 | 0.6815 | 0.8265 | 0.8271 |
| B5 | 0.6812 | 0.6791 | 0.8228 | 0.8235 |
| Avg. | 0.6884 | 0.6837 | 0.8268 | 0.8275 |

Table 3. Comparison of averages

| Technique | Accuracy | F1 Score | AUC | N_MCC |
|-----------|----------|----------|--------|--------|
| MGT_OBWB | 0.7240 | 0.7197 | 0.8466 | 0.8472 |
| MGT_OBHB | 0.6884 | 0.6837 | 0.8268 | 0.8275 |
| OB [24] | 0.5823 | 0.5814 | 0.7669 | 0.7681 |
| OBWB [24] | 0.5849 | 0.5851 | 0.7694 | 0.7695 |
| OBHB [24] | 0.6105 | 0.6112 | 0.7836 | 0.784 |
| MGT_OB [24] | 0.69843 | 0.6918 | 0.8324 | 0.8332 |

Table 4. TL vs VL (Epoch:150)

| | Proposed methods | | | | Without DA | | Methods [24] | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MGT_OBWB | | MGT_OBHB | | OB | | OBWB | | OBHB | | MGT_OB | |
| | TL | VL | TL | VL | TL | VL | TL | VL | TL | VL | TL | VL |
| B1 | 0.04488 | 1.65843 | 0.52446 | 0.90705 | 0.000311 | 4.287225 | 0.000153 | 4.84651 | 0.00093 | 3.69605 | 0.60450 | 0.98230 |
| B2 | 0.04321 | 1.81255 | 0.08188 | 2.03424 | 2.851e-5 | 3.924416 | 5.366e-6 | 4.62421 | 5.68e-6 | 4.87542 | 0.55612 | 0.95762 |
| B3 | 0.04687 | 1.62437 | 0.07377 | 1.79544 | 4.292e-5 | 4.078012 | 3.035e-6 | 5.82033 | 6.701e-6 | 4.97896 | 0.57654 | 0.91903 |
| B4 | 0.04984 | 1.70860 | 0.07997 | 1.77299 | 1.805e-5 | 4.579904 | 5.942e-6 | 4.50964 | 4.878e-6 | 4.63333 | 0.49488 | 0.98593 |
| B5 | 0.04574 | 1.62888 | 0.07005 | 1.79628 | 2.267e-5 | 4.585097 | 5.785e-6 | 4.73261 | 5.520e-6 | 4.84234 | 0.54657 | 0.91448 |
| Avg | 0.04611 | 1.68657 | 0.16603 | 1.66120 | 8.459e-5 | 4.290931 | 3.460e-5 | 4.90666 | 0.000191 | 4.60522 | 0.55572 | 0.95187 |

The graphs for the TL/VL vs epochs are presented for visual interpretations in Figure 2. Analysis of these graphs confirm that the proposed techniques refer Figures 2(a) and 2(b) are better generalized as compared to OB, OBWB, and OBHB techniques Figures 2(c), 2(e), and 2(f). Analysis of the graph for MGT_OB refer Figure 2(d) shows the VL is almost approaching the TL indicating its better generalization ability compared to other techniques.

The generalization gap is the absolute difference between TL and VL. It can be used to find if the model is overfitting and needs more data/better feature dataset for the training. Table 5 summarizes the generalization gap for the methods mentioned in Table 3. It is observed that MGT_OB has the least generalization gap indicating its better generalization capability compared to other methods. OBWB method has the most generalization gap showing that the model is prone to overfitting. It is also observed that for the OB method where experiments are done without augmentation, the generalization ability is better than that of the OBWB and OBHB. This indicates that some IDA methods may lead to overfitting. Note that the performance of OB is poor with respect to all evaluation parameters refer Table 3.

Table 5. Summary of the generalization gap of IDA techniques

| IDA method | TL | VL | |TL-VL| |
|------------|--------|----------|----------|
| MGT_OBWB | 0.04611 | 1.68657 | 1.64046 |
| MGT_OBHB | 0.16603 | 1.66120 | 1.49517 |
| OB [24] | 8.46E-05 | 4.29E+00 | 4.289915 |
| OBWB [24] | 3.46E-05 | 4.91E+00 | 4.909965 |
| OBHB [24] | 0.000191 | 4.60522 | 4.605029 |
| MGT_OB [24] | 0.55572 | 0.95187 | 0.39615 |

## 4.5. Computational complexity

The computational complexity of the proposed methods involve complexity due to GT in addition to complexity due to Hadamard/Walsh. Consider a grayscale image of size nXn. GTs such as shift and flip operations involve copying the pixel values to the desired location and require constant time. A rotation operation on this image requires $n^2*4$ multiplications and $n^2*2$ additions. A fast Hadamard/Walsh transform requires $n\log_2 n$ additions. Thus, the computational complexity of the proposed DA technique when rotation is applied after Hadamard/Walsh is given by in (13).

$$\text{Complexity} = (4 * n^2)\text{Multiplications} + (n^2 + n\log_2 n)\text{Additions} \qquad (13)$$

### 4.6. Limitations of the approach

While applying the proposed methods for medical images datasets values for rotation, and width/height shift. and the % of high energy coefficients are required to be investigated to obain realistic images.
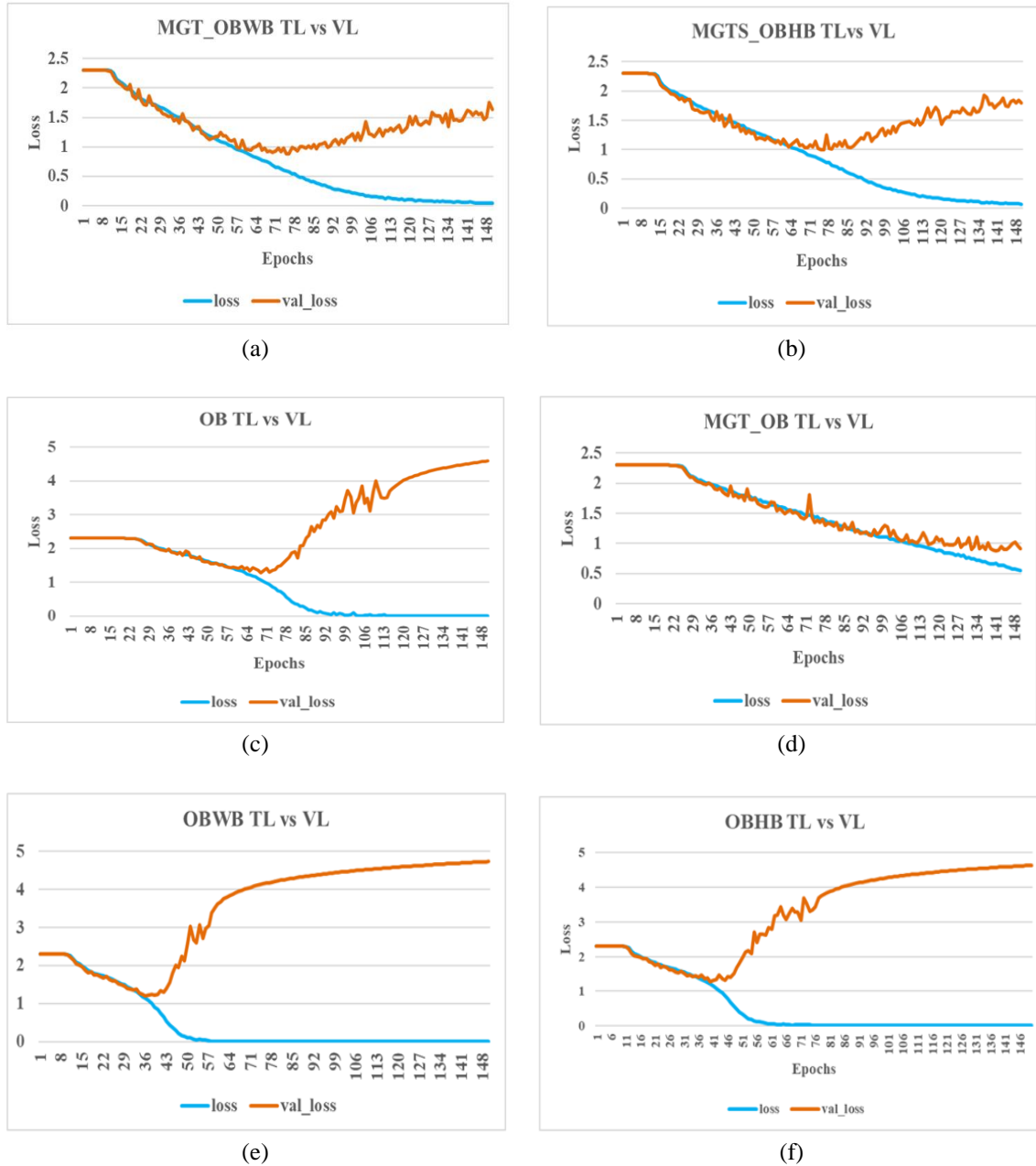


Figure 2. TL vs VL for a dataset containing: (a) images obtained using the proposed MGT_OBWB technique, (b) images obtained using the proposed MGT_OBHB technique, (c) original batch images, (d) original batch images and images obtained using MGT on original batch images, (e) original images and Walsh images, and (f) original images and Hadamard images

### 5. CONCLUSION

In this research, two hybrid methods using frequency domain coefficients (obtained using HT and Walsh transform) and geometrical transformations is developed for IDA. The proposed methods are used to

create an enormous dataset of size 100000 from each batch of size 10000 images. The performance of the proposed IDA methods on VGG-16 model has shown significant improvement. Further, analysis shows that the proposed methods are better in addressing the overfitting compared to other methods. We also gain an insight from our empirical study that some IDA methods are not very effective in addressing the overfitting issue and the researchers must do thorough analysis before finalizing the IDA method.

## REFERENCES

[1]     A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
[2]     Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
[3]     C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.
[4]     L. Taylor and G. Nitschke, "Improving deep learning with generic data augmentation," in *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, IEEE, Nov. 2018, pp. 1542–1547, doi: 10.1109/SSCI.2018.8628742.
[5]     M. Elgendi *et al.*, "The effectiveness of image augmentation in deep learning networks for detecting COVID-19: a geometric transformation perspective," *Frontiers in Medicine*, vol. 8, Mar. 2021, doi: 10.3389/fmed.2021.629134.
[6]     D. Ghosh, "Improved COVID-19 detection using data augmentation deep convolution GAN and classiier DenseNet," *Research Square,* 2021, doi: 10.21203/rs.3.rs-235624/v1.
[7]     Y. Hou and M. Navarro-Cía, "A computationally-inexpensive strategy in CT image data augmentation for robust deep learning classification in the early stages of an outbreak," *Biomedical Physics & Engineering Express*, vol. 9, no. 5, Sep. 2023, doi: 10.1088/2057-1976/ace4cf.
[8]     M. Nishio, S. Noguchi, H. Matsuo, and T. Murakami, "Automatic classification between COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy on chest X-ray image: combination of data augmentation methods," *Scientific Reports*, vol. 10, no. 1, Oct. 2020, doi: 10.1038/s41598-020-74539-2.
[9]     A. Azade and K. M. Anand, "Impact of image augmentation in COVID-19 detection using chest X-ray images," in *2022 IEEE Delhi Section Conference (DELCON)*, IEEE, Feb. 2022, pp. 1–5, doi: 10.1109/DELCON54057.2022.9752785.
[10]    N. Yella and B. Rajan, "Data augmentation using GAN for sound based COVID 19 diagnosis," in *2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, IEEE, Sep. 2021, pp. 606–609, doi: 10.1109/IDAACS53288.2021.9660990.
[11]    R. Hu, G. Ruan, S. Xiang, M. Huang, Q. Liang, and J. Li, "Automated diagnosis of COVID-19 using deep learning and data augmentation on chest CT," *medRxiv*, 2020, doi: 10.1101/2020.04.24.20078998.
[12]    S. Motamed, P. Rogalla, and F. Khalvati, "Data augmentation using generative adversarial networks (GANs) for GAN-based detection of Pneumonia and COVID-19 in chest X-ray images," *Informatics in Medicine Unlocked*, vol. 27, 2021, doi: 10.1016/j.imu.2021.100779.
[13]    E. C. Adjobo, A. T. S. Mahama, P. Gouton, and J. Tossa, "Towards accurate skin lesion classification across all skin categories using a PCNN fusion-based data augmentation approach," *Computers*, vol. 11, no. 3, Mar. 2022, doi: 10.3390/computers11030044.
[14]    F. Perez, C. Vasconcelos, S. Avila, and E. Valle, "Data augmentation for skin lesion analysis," in *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*, 2018, pp. 303–311, doi: 10.1007/978-3-030-01201-4_33.
[15]    T. C. Pham, C. M. Luong, M. Visani, and V. D. Hoang, "Deep CNN and data augmentation for skin lesion classification," in *Intelligent Information and Database Systems (ACIIDS 2018)*, 2018, pp. 573–582, doi: 10.1007/978-3-319-75420-8_54.
[16]    H. Rashid, M. A. Tanveer, and H. Aqeel Khan, "Skin lesion classification using GAN based data augmentation," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),* Berlin, Germany, 2019, pp. 916-919, doi: 10.1109/EMBC.2019.8857905.
[17]    A. Saha, P. Prasad, and A. Thabit, "Leveraging adaptive color augmentation in convolutional neural networks for deep skin lesion segmentation," in *International Symposium on Biomedical Imaging*, IEEE, Apr. 2020, pp. 2014–2017, doi: 10.1109/ISBI45749.2020.9098344.
[18]    N. R. Swetha, V. K. Shrivastava, and K. Parvathi, "Multiclass skin lesion classification using image augmentation technique and transfer learning models," *International Journal of Intelligent Unmanned Systems*, vol. 12, no. 2, pp. 220–228, Jun. 2021, doi:10.1108/ijius-02-2021-0010.
[19]    S. Shen *et al.*, "A low-cost high-performance data augmentation for deep learning-based skin lesion classification," *BME Frontiers*, vol. 2022, Jan. 2022, doi: 10.34133/2022/9765307.
[20]    Y. Shen, J. Gu, X. Tang, and B. Zhou, "Interpreting the latent space of GANs for semantic face editing," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE, Jun. 2020, pp. 9240–9249, doi: 10.1109/CVPR42600.2020.00926.
[21]    L. Fetty *et al.*, "Latent space manipulation for high-resolution medical image synthesis via the StyleGAN," *Zeitschrift fur Medizinische Physik*, vol. 30, no. 4, pp. 305–314, Nov. 2020, doi: 10.1016/j.zemedi.2020.05.001.
[22]    A. R. Luaibi, T. M. Salman, and A. H. Miry, "Detection of citrus leaf diseases using a deep learning technique," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 2, pp. 1719–1727, Apr. 2021, doi: 10.11591/ijece.v11i2.pp1719-1727.
[23]    K. Meethongjan, V. T. Hoang, and T. Surinwarangkoon, "Data augmentation by combining feature selection and color features for image classification," *International Journal of Electrical and Computer Engineering*, vol. 12, no. 6, pp. 6172–6177, Dec. 2022, doi: 10.11591/ijece.v12i6.pp6172-6177.
[24]    V. Suryawanshi, T. Sarode, N. Jhunjhunwala, and H. Khan, "Evaluating image data augmentation technique utilizing hadamard Walsh space for image classification," *Proceedings of International Conference on Intelligent Vision and Computing (ICIVC 2022),*

2023, pp. 290–301, doi: 10.1007/978-3-031-31164-2_24.

[25]   V. Suryawanshi, S. Adivarekar, K. Bajaj, and R. Badami, "Comparative study of regularization techniques for VGG16, VGG19 and ResNet-50 for plant disease detection," *Proceedings of International Conference on Communication and Computational Technologies (ICCCT 2023)*, 2023, pp. 771–781, doi: 10.1007/978-981-99-3485-0_61.

[26]   A. Krizhevsky, V. Nair, and G. Hinton, "CIFAR-10 and CIFAR-100 datasets," *University of Toronto*, 2009. [Online]. Available: https://www.cs.toronto.edu/~kriz/cifar.html

[27]   A. K. Jain, *Fundamentals of digital image processing*, New Delhi: Prentice Hall of India, 2006.

[28]   K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, Sep. 2015.

[29]   D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, Dec. 2020, doi: 10.1186/s12864-019-6413-7.

[30]   I. Markoulidakis, I. Rallis, I. Georgoulas, G. Kopsiaftis, A. Doulamis, and N. Doulamis, "Multiclass confusion matrix reduction method and its application on Net Promoter Score classification problem," *Technologies*, vol. 9, no. 4, Nov. 2021, doi: 10.3390/technologies9040081.

# BIOGRAPHIES OF AUTHORS

**Vaishali Suryawanshi** ⓘ 🔱 SC ▷ holds a is an Assistant Professor at Thadomal Shahani Engineering College, Mumbai. She received her B.E. in Computer Engineering from the North Maharashtra University, Jalgaon. She completed her M.E. (Computer Engineering) from the University of Mumbai. She is presently pursuing her Ph.D. in Computer Engineering from University of Mumbai. She has published 15 research articles in reputed journals and conferences. She has received best paper award for two research articles. She is a lifetime memebr of ISTE and CSI. Her area of interest are deep learning, machine learning, artificial intelligence, and computer vision. She can be contacted at email: vaishali.suryawanshi@thadomal.org.

**Dr. Tanuja K. Sarode** ⓘ 🔱 SC ▷ has received Ph.D. degree from Mukesh Patel School of Technology, Management and Engg., Mumbai, and M.E. (Computer Engg.) degree from Mumbai University. She has more than 23 years of experience in teaching. Currently working as a Professor in the Dept. of Computer Engg. at Thadomal Shahani Engineering College, Mumbai. She is Life member of ISTE and IETE. Her areas of interest are deep learning, machine learning, aritifical intelligence, image processing, signal processing and computer graphics. She has more than 200 papers in International Conferences or journal to her credit. She can be contacted at email: tanuja.sarode@thadomal.org.