Sriwijaya Electrical and Computer Engineering Journal (SELCO)
Vol. 1, No. 2, September 2024, pp. 44-53
DOI: 10.62420/selco.v1i2.6

Image processing using sparse representation for classification with semi-random projection dimension reduction for the image recognition system

Izzatul Jannah [1], Puspa Kurniasari [1]
[1] Department of Electrical Engineering, Universitas Sriwijaya, Jl. Raya Palembang-Prabumulih KM 32, Indralaya 30662, Indonesia

ARTICLE INFO

Article history:

Received October 7, 2024
Revised November 12, 2024
Published December 18, 2024

Keywords:

Sparse representation;
Classification;
Semi random projection;
GUI;
PyCharm;
PSNR

ABSTRACT

Image recognition is a technology to identify objects, places, people and several other variables in digital images. The algorithm that can be used in the image recognition system is Sparse Representation for Classification. However, the high computational load is a problem in this study. In addition, there is a lot of training data needed to meet the sparse condition which is a weakness of this algorithm. Thus, to overcome this problem, dimensionality reduction can be carried out on the image using the Semi Random Projection method. In this study, the author used PyCharm software to process the Guide User Interface (GUI) of the dimensionality reduction system and image recognition using Semi Random Projection-Sparse Representation for Classification. Testing was carried out using 100 training data in the form of 50 red-green-blue images and 50 grayscale images. The images are divided into 10 classes and use 10 test images that have been added with noise and occlusion. Testing was carried out five times each on each test image. From the testing in this study, the results obtained on good performance parameters with an average accuracy of 98.93%, an average Peak Signal Noise to Ratio (PSNR) of 33.392741 dB and an average computing time of 1112.57952 ms.

Corresponding Author:

Puspa Kurniasari, Department of Electrical Engineering, Universitas Sriwijaya, Jl. Raya Palembang-Prabumulih KM 32, Indralaya 30662, Indonesia
Email: puspakurniasari@ft.unsri.ac.id

1. INTRODUCTION

Signal is a form of data representation that contains information with physical quantities that change in space, time, or other independent variables. In order for the signal to be sent properly, the signal will go through a processing process. Signal processing in machine learning can be used for various implementations such as biometric recognition and image recognition [1]. Image recognition is a technology for identifying objects, places, people, and several other variables in digital images. There are three stages in image recognition, namely image detection, feature extraction, and classification [2]. The algorithm that can be used to implement image recognition is Sparse Representation for Classification [3]. However, the high computational load and the need for a lot of training data to meet the sparse conditions are weaknesses of this algorithm. Thus, image dimensionality reduction can be done to overcome this problem. Image dimensionality reduction can be done using the Semi-Random Projection (SRP) method. This method is a development of the Random Projection (RP) method. Random Projection states that if the points in the vector space have a high enough dimension, then the points can be projected into the appropriate low-dimensional space, while maintaining the distance between the points. Thus, RP will help speed up the feature extraction process [4]. However, the difference between the SRP method and the RP method is that the SRP transformation vector is obtained from studying the data while the RP transformation vector is generated randomly.

Semi-random projection aims to find a latent space with high discriminatory power but still has a reasonable computational burden on the image recognition system [5]. Image processing using Sparse Representation for Classification (SRC) on the reduced dimension image using Semi-Random Projection for the image recognition system in this study. Therefore, this proposed image processing research is not only based on test images with facial objects but also involves occlusion variations in the test image. Thus, this study proposes an image recognition system where the image classification used in processing is broader, including Red-Green-Blue (RGB) images, Grayscale images, and various objects and personal datasets for public images and Yale datasets for facial images, then there is one system derived from the combination of two algorithms used, namely Semi-Random Projection for the reduction system and Sparse Representation for Classification for the image recognition system. These two systems are interrelated because images that have not been processed using dimension reduction cannot be processed into the recognition system, then the results obtained are analyzed by considering performance parameters, namely Peak Signal Noise Ratio (PSNR), sparse matrix value, computing time, accuracy.

2. BASIC THEORY

2.1. Sparse Representation for Classification

The Sparse Representation for Classification algorithm is an algorithm proposed by Wright, et al. in 2009.The main idea of this algorithm is to represent the test image as a sparse linear combination of all the training images [3]. Sparse representation can perform quite good face classification, even though there is significant noise and occlusion. SRC has been widely used in image processing and non-linear differential equations.

2.2. L_1 Minimization

L_1 Minimization is a method that is quite efficient in completing signal recovery with sparse conditions in acertain linear equation. This method can be applied in various fields and has been proven effective, such as geophysics, image processing, and data compression [3].

2.3. Semi-Random Projection

Semi-Random Projection is a development of the Random Projection method. The core idea behind random projection is that if the points in a vector space have high enough dimensions then they can be projected into a suitable low-dimensional space while maintaining the distance between the points. SRP consists of two phases, namely random sampling of features and learning of transformation vectors. It is called Semi-Random because the first phase is random while the second is not. This phase is what differentiates between RP and SRP. The transformation vector in RP will be chosen randomly while the SRP transformation vector is obtained from studying the data [5].

2.4. Peak Signal To Noise Ratio

Peak Signal to Noise Ratio is a comparison between the maximum value of the measured image bit depth (8-bit image, has a maximum value of 255) and the amount of noise that affects the signal. PSNR is usually measured in decibels. PSNR is used to determine the comparison of image quality before and image after processing. The minimum PSNR value limit is generally 30 dB. The greater the PSNR value, the better the image processing results or closer to the original image [6].

2.5. Sparse Matrix

A sparse matrix is a matrix that consists mostly of zero values. The sparseness of a matrix result can be quantified with a score which is the number of zero values in the matrix divided by the total number of elements in the matrix. The sparse matrix value will determine the class where the test data belongs [7].

2.6. Accuracy

Accuracy in machine learning is a representation of the overall truth of a model's predictions. Accuracy calculations can be obtained by subtracting the full accuracy value from the error value [8].

2.7. Computing Time

Computing time is the time used to carry out testing or processing on a system. The less computing time required, the better the system used [9].

3. METHODS

3.1. Hardware and Software Preparation

In this research, there are several hardware devices used as media in processing programming and retrieving training data as well as several software used to carry out programming functions which are listed in Table 1.

Table 1. Hardware and Software

No.	Device Name		Device
	Hardware	Software	Specifications
1.	HP Laptop 14-bs1xx		*Windows* 10 Intel® *Core*™ i58250U CPU 4096 MB RAM, 914 GB HDD
2.	*iPhone 6s Plus*		IOS 14 12 MP, dual LED flash 64 GB + RAM 2 GB
3.		PyCharm	2023
4.		Google Colab	Free
5.		Canva	Free

3.2. Research Flow Diagram

Based on Fig. 1 about the research flow diagram that the author has compiled, the process begins by preparing training data images. The training data images consist of 50 RGB images and 50 Grayscale images with original dimensions and have been divided into ten different classes.

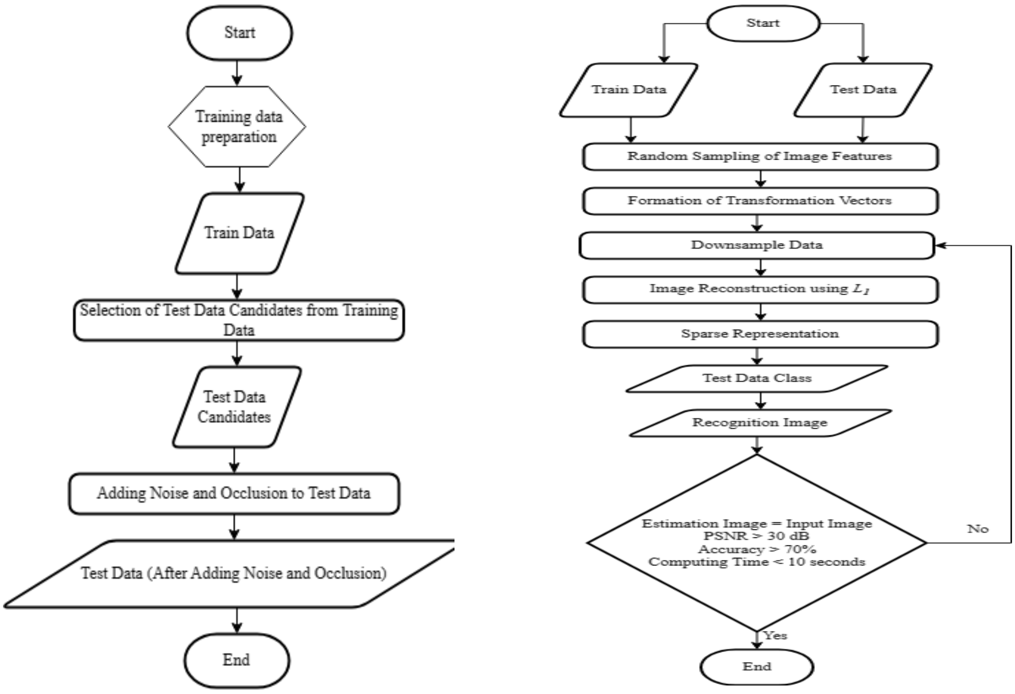

Fig. 1. Research flow diagram

These images are a private data set for the general image class, namely book images, flower images, cat images, and mosque images, and use the Yale Database for the facial image class. The reason for using the Yale Database is because this data set is widely used for facial recognition research which consists of 15 individuals with eleven images each with various conditions such as different lighting factors, poses, and expressions. Based on the training data, the stage of selecting candidate test data images is continued. These images are ten test images from ten classes of training data, meaning one image from one class is taken as a candidate for test data, and then the ten candidate test data enter the addition of noise and occlusion. For general images, noise is added on the form of Salt and Pepper Noise, while for facial images, occlusion is added in the form of fake moustaches and glasses. These ten images will then become test data, then on each image data a random sampling of features will be carried out from all the original features and the image data will be extracted from the features taken to form a transformation vector and then continued with the downsampling process then a sparse representation search will be carried out using matrix reconstruction and then obtain the sparse matrix coefficients which indicate the class the test data is in so that the recognition image is obtained. If the image resulting from recognition is the same as the input image, the PSNR comparison of the original image and the image resulting from recognition is above a minimum of 30 db, the accuracy is above 70% and the computing time is below 10 seconds then the process is complete. However, if the image resulting from the recognition is different from the input image, the PSNR comparison of the original image and the image resulting from the recognition is below a minimum of 30 db, the accuracy is below 70% and the computing time is above 10 seconds, then the process returns to downsampling the data.

4. RESULTS AND DISCUSSION

In this section, the research results obtained show that the system has succeeded in producing estimated image results through sparse representation for classification with semi-random projection reduction. The stages applied in this study to produce system output in the form of estimated images are determining training data and test data from image classes, random sampling, image downsampling, image reconstruction, and testing of image output from the method used in the system.

4.1. Determination of Training Data

The training data used are ten image classes divided based on image color, namely five Red Green Blue (RGB) classes and five Grayscale (GS) classes. The ten classes include the RGB Book class, RGB Flower class, RGB Cat class, RGB Subject class, RGB Mosque class, Grayscale Book class, Grayscale Flower class, Grayscale Cat class, Grayscale Subject class and Grayscale Mosque class. Each class contains ten images so that the total number of training data images is 100 images. The training data for one of the RGB classes can be seen in Fig. 2 and the grayscale class can be seen in Fig. 3. This original image training data except for the subject class image was taken directly using a smartphone camera in indoor and in outdoor locations during the day and evening around at 11 a.m to 5 p.m with natural lighting, namely the sun when outdoors and indoor lighting assistance in the form of lamps and image capture angles, namely front view, side view, top view and bottom view. Therefore, for the Book Class, Flower Class, Cat Class, and Mosque Class, the images are a collection of personal data taken using a smartphone camera with different lighting conditions, angles, times, and places.

Fig. 2. RGB book class

Fig. 3. Grayscale Flower Class

As for the Subject Class, the image is the official Yale dataset taken from the internet, namely through the site kaggle.com/datasets/olgabelitskaya/yale-face-database. Because the original color of the Book Class, Flower Class, Cat Class, and Mosque Class images is RGB, a color conversion is carried out from RGB to grayscale to obtain the Grayscale class and for the Subject Class, a color conversion is carried out from grayscale to RGB to obtain the RGB class because the original color of the Subject image is Grayscale. This color conversion process is carried out online through the site onlinetools.com/image/grayscale-image.

4.2. Determination of Test Data

Noise can occur due to sharp interference in the image signal so in the test data of books, flowers, cats, and mosques, the process of noise adding in the form of Salt and Pepper Noise was carried out. The Salt and Pepper Noise is noise in the form of black and white dots in an image like a scattering of salt and pepper. This noise was chosen because it has a denser noise level than other types of noise. The Salt and Pepper Noise addition process was carried out using the Google Colab application and Python programming. There are subject test data because the image is a facial image, the process of occlusion adding was carried out in the form of a fake moustache and glasses which are representations of properties that humans often wear on their faces in real conditions. This occlusion addition process was carried out using the Canva application. The purpose of noise adding and the occlusion adding is to disguise the image so that it has a difference compared to the original image. This noise adding and occlusion are also representations of disturbances that often occur in real conditions. The overall results of the images that became the test data after the process of adding noise and occlusion can be seen in Fig. 4.

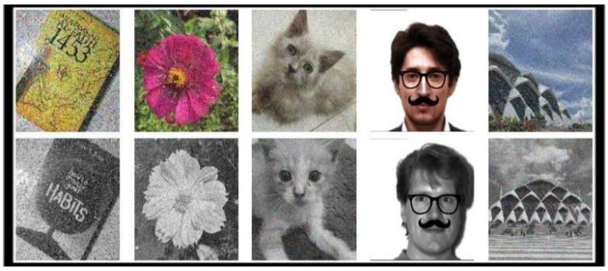

Fig. 4. Test Data

4.3. Random Sampling of Image Features

In the system, one of the selected test data is first subjected to a dimension reduction process, then the image recognition process is performed on the test data. This entire process is carried out using the PyCharm application with the Python and Sklearn programming languages. The Sklearn library is specifically used for the Semi-Random Projection dimension reduction process. This sampling is carried out on training data and test data. The features contained in the training data image and the test data image are sampled randomly. For RGB images, the color channels are divided into r (red), g (green), and b (blue). This channel separation is carried out so that sampling is more focused on each color channel. The color channels are shown in Fig. 5 while Fig. 6 is the projection data from the RGB image. Fig. 6 shows the red channel marked with a red dot, the green channel marked with a green dot, and the blue channel marked with a blue dot so that a randomly selected feature sample is obtained from the projection data marked with a dark green dot.

Sriwijaya Electrical and Computer Engineering Journal (SELCO)
Vol. 1, No. 2, September 2024, pp. 44-53
DOI: 10.62420/selco.v1i2.6

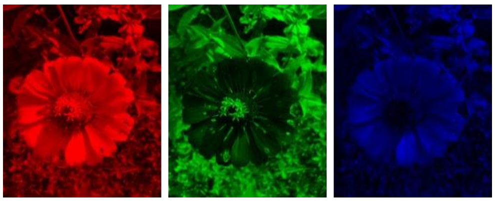

Fig. 5. The Color Channels of Red-Green-Blue

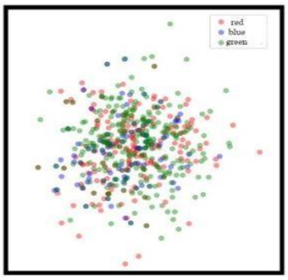

Fig. 6. RGB Image Projection Data

Fig. 7 is projection data from a grayscale image which shows the distribution of data and samples selected by the algorithm. In Grayscale images, channel splitting is not performed and the data is directly projected. The data projection is marked with a light gray dot and a randomly selected feature sample is obtained from the projection data which is marked with a dark gray dot.

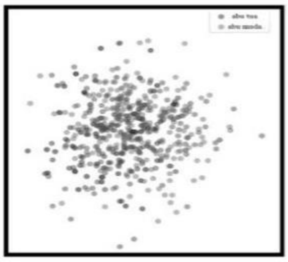

Fig. 7. Grayscale Image Projection Data

4.4. Formation of Transformation Vectors

From the random sample of image features based on Fig. 6 and Fig. 7, the image data is then extracted from the selected features to form a transformation vector so that the data is reshaped into a single-column vector shown in Fig. 8(a) and Fig. 8(b). After that, the system obtains an image that has been reduced in dimension using the dimension reduction method and then the image is processed in the recognition process.

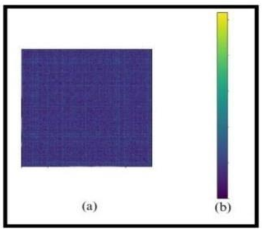

Fig. 8.(a). Data Extract **(b).** Vector Transformation

4.5. Image Downsample

The reduced data is then downsampled. Downsampling aims to allow the image matrix to perform column normalization so that image feature extraction is obtained as in Fig. 9 is the downsample result of the RGB test image and it can be seen that in the image when the RGB image is downsampled, the resolution decreases and it has a pixelated appearance and only the color and shape of the image are visible.

Fig. 9. Downsample RGB Image

Grayscale image downsample can be seen in Fig. 10 as the downsample result of the grayscale test image which shows that the downsample process causes the image to experience a decrease in image resolution and only the shape of the image is also visible.

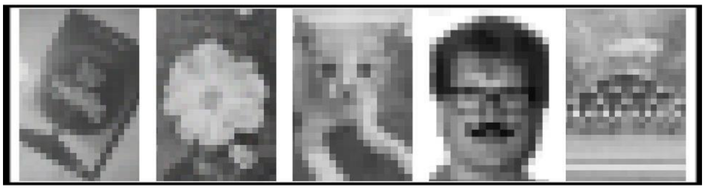

Fig. 10. Downsample Grayscale Image

4.6. Image Reconstruction Using L_1 Minimization

The downsampled image is then reconstructed through L_1 Minimization into the original image, eliminating noise and occlusion in the image. The results of the RGB image reconstruction can be seen in Fig. 11. Fig. 11 is the result of the RGB test image reconstruction and it can be seen that in the image when the RGB image reconstruction was carried out, the noise was removed. The results of the Grayscale image reconstruction can be seen in Fig. 12 which shows the changes of the test image and noise removal so that the test image is getting closer to the original image.

Fig. 11. RGB Image Reconstruction Results

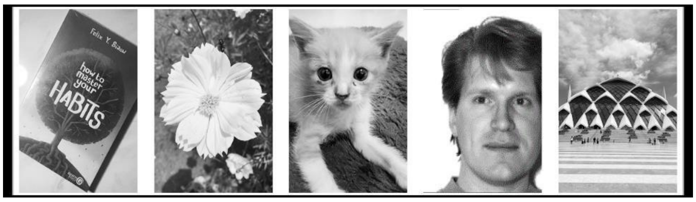

Fig. 12. Grayscale Image Reconstruction Results

4.7. Sparse Representation for Classification

Sparse representation is performed on the reconstructed image to obtain the sparse matrix value for all training data using the sparse coefficient. The sparse coefficient is obtained by multiplying the inverse training matrix by the test matrix. This sparse matrix value then determines the location of the test data class as shown in Fig. 13. The test data is in the class whose image has the largest sparse matrix value as shown in Fig. 13 as the result of testing one of the RGB images, namely RGB Book. This sparse matrix value also indicates the condition of the error value obtained. In Fig. 13, the training data class 0 to 10 is the Book class, class 10 to 20 is the Flower class, class 20 to 30 is the Cat class, class 30 to 40 is the Subject class and 40 to 50 is the Mosque class. Based on Fig. 13, the largest sparse matrix value, namely 0.763, is from the Book class so it is known that the book test image is in the Book class and the error is obtained from all training data. The Book class indicates the lowest error at 0.869%. Thus, the resulting image recognition result is obtained with a computing time of 633.3204 ms which means that in conducting image testing, the time required is 0.633 seconds.

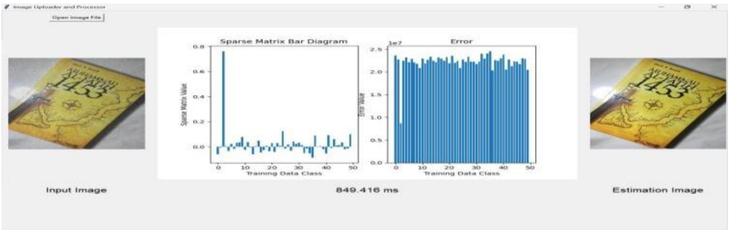

Fig. 13. Test Result

The sparse matrix value of the image test in **Fig. 13** using RGB Book can be seen in **Fig. 14**

1	-6.11E-02	26	-1.32E-02
2	-6.25E-03	27	1.83E-02
3	7.63E-01	28	-3.15E-02
4	5.32E-03	29	4.21E-02
5	-3.24E-02	30	2.18E-02
6	2.21E-02	31	3.29E-02
7	-1.56E-02	32	1.33E-02
8	3.31E-02	33	-4.98E-02
9	3.49E-02	34	-1.37E-02
10	7.76E-02	35	-5.24E-02
11	-2.44E-02	36	-8.55E-02
12	3.62E-02	37	9.03E-02
13	-4.03E-03	38	-9.63E-04
14	-5.94E-02	39	4.87E-03
15	4.99E-03	40	-1.82E-02
16	4.94E-02	41	-5.33E-02
17	-4.55E-02	42	9.23E-02
18	-2.50E-02	43	-8.17E-03
19	8.58E-03	44	6.05E-02
20	-3.59E-02	45	1.06E-02
21	2.95E-02	46	1.20E-02
22	-4.07E-02	47	3.49E-02
23	2.93E-02	48	-1.96E-02
24	8.15E-03	49	-1.43E-02
25	1.23E-01	50	9.98E-02

Fig. 14. Sparse Matrix Result

From Fig. 14, it can be seen that the rgbbuku03.jpg image in class number 3 with 0.763 is the largest sparse matrix value compared to other classes. Thus, the error obtained from the rgbbuku03_uji.jpg image is the lowest, which is 0.869% as in Fig. 15.

1	2.36E+00	26	2.27E+00
2	2.27E+00	27	2.22E+00
3	8.69E-01	28	2.33E+00
4	2.25E+00	29	2.22E+00
5	2.32E+00	30	2.23E+00
6	2.21E+00	31	2.17E+00
7	2.29E+00	32	2.22E+00
8	2.20E+00	33	2.40E+00
9	2.18E+00	34	2.29E+00
10	2.09E+00	35	2.41E+00
11	2.29E+00	36	2.45E+00
12	2.18E+00	37	2.03E+00
13	2.26E+00	38	2.26E+00
14	2.33E+00	39	2.24E+00
15	2.25E+00	40	2.30E+00
16	2.21E+00	41	2.38E+00
17	2.32E+00	42	2.05E+00
18	2.30E+00	43	2.27E+00
19	2.24E+00	44	2.13E+00
20	2.33E+00	45	2.24E+00
21	2.20E+00	46	2.23E+00
22	2.35E+00	47	2.17E+00
23	2.20E+00	48	2.30E+00
24	2.24E+00	49	2.29E+00
25	2.08E+00	50	2.05E+00

Fig. 15. Error Result

4.8. Test Results of Image Recognition-Dimension Reduction System using Semi Random Projection and Sparse Representation for Classification

In testing the Semi-Random Projection-Sparse Representation for Classification system using PyCharm software, this system is displayed in the form of a Guide User Interface (GUI). The testing process on each test image was carried out five times. This aims to obtain patterns and comparisons between the first test and other tests based on test parameters. Image test results can be seen in Table 2.

Table 2. Image Testing Results

Image	Sparse Matrix Value	Error (%)	Average Computing Time(ms)	Average PSNR (dB)	Accuracy (%)
Book RGB	0.763	0.869	622.81	33.125	99.131
Flower RGB	0.749	0.726	506.32	32.345	99.274
Cat RGB	0.762	0.97	2100.691	34.46	99.03
Subject RGB	0.987	1.61	303.391	39.41	98.39
Mosque RGB	0.75	0.9	1972.213	31.98	99.1
Book Grayscale	0.821	1.06	572.141	32.57	98.94
Flower Grayscale	0.884	1.08	472.005	33.02	98.92
Cat Grayscale	0.942	0.87	2136.622	31.5	99.13
Subject Grayscale	0.951	1.57	322.711	33.88	98.43
Mosque Grayscale	0.815	1.04	2116.89	31.66	98.96
Average	**0.8424**	**1.07**	**1112.579**	**33.395**	**98.93**

Based on Table 2, it can be seen that from all tests the sparse matrix and error values obtained are constant, meaning that changes in sparse matrix values and error values in the SRP-SRC system are only influenced by the input image. The average computing time obtained was 1112.579 ms or 1.112 seconds. Based on the minimum standard PSNR value of 30 dB, the average PSNR value obtained is 33.395 dB, indicating that the recognition image has good image quality. The accuracy obtained is 98.93% so it can be seen that with this average accuracy, image recognition using Sparse Representation for Classification with dimension reduction using Semi-Random Projection has good performance, meaning that every test image contains noise and occlusion in it and has gone through the process. The dimensional reduction both RGB and grayscale can be still recognized well by the system.

5. CONCLUSION

Based on tests carried out by the author, it can be concluded that the Semi-Random Projection-Sparse Representation for the Classification algorithm has succeeded in reducing image dimensions and recognizing images with an average accuracy value of 98.93%, meaning that every test image contains noise and occlusion in it and has through the dimension reduction process, both RGB and Grayscale (GS) can still be recognized well by the system and based on the minimum standard PSNR value of 30 dB, the average PSNR value obtained is 33,395 dB, indicating that the image resulting from the recognition has good image quality. In future research, the use of linear dimension reduction methods such as Principal Component Analysis to obtain better accuracy values. The addition of the number of training data classes and test data with more diverse images also needs to be done including RGB, Grayscale, and Binary so that it can be known how well the system can work. In addition, the use of moving objects such as videos or gifs can be used in research development to expand the implementation of dimension reduction algorithms and image recognition algorithms.

REFERENCES

[1] A. Sharma, S. Arya and P. Chaturvedi, "A Novel Image Compression Based Method for Multispectral Fingerprint Biometric System," *Third International Conference on Computing and Network Communications (CoCoNet'19)*, *Procedia Computer Science*, vol. 171, pp. 1698-1707, 2020, https://doi.org/10.1016/j.procs.2020.04.182.

[2] V. Tyagi, "Understanding Digital Image Processing," 1st ed. *CRC Press*, 2018, doi: 10.1201/9781315123905.

[3] Z. Zha, X. Yuan, B. Wen, J. Zhang, J. Zhou and C. Zhu, "Image Restoration Using Joint-Patch-Group Based Sparse Representation," *IEEE Transactions on Image Processing*, vol. 29, pp. 7735-7750, 2020, doi: 10.1109/TIP.2020.3005515.

[4] H. Xie, J. Li, and H. Xue, "A survey of dimensionality reduction techniques based on random projection," June. 2017, pp. 1-11. [Online]. Available: http://arxiv.org/abs/1706.04371.

[5] R. Zhao and K. Mao, "Semi-Random Projection for Dimensionality Reduction and Extreme Learning Machine in High-Dimensional Space," *IEEE Computational Intelligence Magazine*, vol. 10, issue: 3, pp. 30–41, Aug. 2015, doi: 10.1109/MCI.2015.2437316.

[6] S. M. Belgaonkar and V. Singh, "Image Compression and Reconstruction in Compressive Sensing Paradigm," *Global Transitions Proceedings*, vol. 3, issue: 1, pp. 220-224, 2022, https://doi.org/10.1016/j.gltp.2022.03.026.

[7] X. Han and H. N. Najm, "Effective Construction of Eigenvectors for A Class of Singular Sparse Matrices," *Applied Mathematics Letters*", vol. 97, November 2019, pp. 121-126, 2019, https://doi.org/10.1016/j.aml.2019.05.018.

[8] KB. Pranav and J. Manikandan, "Design and Evaluation of a Real-Time Face Recognition System using Convolutional Neural Networks," *Third International Conference on Computing and Network Communications (CoCoNet'19)*, *Procedia Computer Science*, vol. 171, pp. 1651-1659, 2020, https://doi.org/10.1016/j.procs.2020.04.177.

[9] J. S. Nayak and M. Indiramma, "An Approach to Enhance Age Invariant Face Recognition Performance Based On Gender Classification," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, issue: 8, part A, September 2022, pp. 5183-5191, 2022, https://doi.org/10.1016/j.jksuci.2021.01.005.

BIOGRAPHY OF AUTHORS

Izzatul Jannah was born in Palembang, South Sumatera, Indonesia in 2001. She received a bachelor's degree from the Department of Electrical Engineering, Universitas Sriwijaya, Indonesia in 2024. Her research was about image processing. She can be contacted at izzatuljannah65@gmail.com

Puspa Kurniasari was received the B.S degree in Electrical Engineering from Universitas Sriwijaya, Indonesia in 2006 and Master Degree from Institute Teknologi Bandung. She is currently lecturer in Department of Electrical Engineering of Universitas Sriwijaya, Indonesia. Her research interest include signal processing and wireless communication. Email : puspakurniasari@ft.unsri.ac.id