**EAI Endorsed Transactions**
**on Internet of Things**

Research Article **EAI.EU**

# Robustness of NMF Algorithms Under Different Noises

Mengyao KANG[1,*], Jiawei ZHAO[1] and Zheng HAN[1]

[1] Faculty of Engineering, The University of Sydney, Sydney, Australia

## Abstract

In machine learning, datasets are often disturbed by different noises. The Nonnegative Matrix Factorization (NMF) algorithm provides a robust method to deal with noise, which will significantly improve the efficiency of machine learning. In this investigation, the standard NMF algorithm and $L_{2,1}$-Norm Based NMF algorithm are studied by designing experiments on different noise types, noise levels, and datasets. Furthermore, Relative Reconstruction Errors (RRE), accuracy, and Normalized Mutual Information (NMI) are used to evaluate the robustness of the two algorithms. In this experiment, there is no significant difference in performance between the two algorithms, while $L_{2,1}$-Norm Based NMF algorithm shows relatively small advantages.

**Keywords:** Machine learning, Nonnegative matrix factorization, Robustness of algorithm, NMF

Received on 21 April 2023, accepted on 21 April 2023, published on 22 June 2023



doi: 10.4108/eetiot.v9i1.3271

## 1. Introduction

Lee and Seung proposed NMF as a method to find matrix factors with the partial global explanation. This algorithm has a certain sparseness when describing data, which makes the data still show relatively robustness when noise and outliers are generated due to external factors (Seung & Lee, 1999). Due to these properties, NMF is widely studied in machine learning. Image analysis and processing is an essential field of machine learning. In image clustering tasks, developers often have to face many images contaminated by noise and outliers. NMF provides a relatively robust processing scheme for such tasks. It uses the matrix decomposition method to divide the image into two matrices and continuously updates by using the critical features extracted from the image. NMF reconstructs the image through iteration to make the new image more similar to the original image to achieve a better clustering effect. In this study, a clustering experiment was designed to add two different noises to two image datasets, which two different NMF algorithms were selected to reconstruct the contaminated images. The reconstructed images were used to conduct the K-Means clustering experiment. Three metrics were used to evaluate the clustering effect, including Relative Reconstruction Errors (RRE), Accuracy, and Normalized Mutual Information (NMI). The study aims to understand the different levels of the robustness of different NMF algorithms under different noise disturbances. It could help select more accurate and stable NMF algorithms in specific machine learning tasks.

## 2. Previous Work

The relevant issues have been explored in past studies by Deguang Kong et al. They considered the influence of two different kinds of noise on the image, including Gaussian noise based on the zero-mean normal distribution and Laplacian noise based on zero-mean Laplacian distribution (Kong et al., 2011). Deguang Kong et al. judged the difference between the standard NMF and the $L_{2,1}$-Norm Based NMF by three criteria of accuracy, NMI, and cluster purity with ten image datasets, including the Yale dataset for testing. The results showed that $L_{2,1}$-Norm Based NMF showed better performance (Kong et al., 2011).

In order to further explore the difference in robustness between the two algorithms, the ORL dataset is added in this

*Mengyao KANG. Email: xiao7sky@outlook.com

experiment. Two different noises are created. Moreover, it also additionally introduces RRE as an evaluation index to explore the robustness of standard NMF and $L_{2,1}$-Norm Based NMF from different perspectives.

## 3. Methods

### 3.1. Standard NMF Algorithm

The objective of NMF is to divide a large matrix into two smaller matrices, which can then be multiplied jointly to form the initial large matrix. The term "non-negative" indicates that there are no negative numbers in the matrix. In other words, all values in the NMF algorithm are positive.

#### Loss Function

The loss function is to measure the quality of the model prediction. In the NMF matrix, the loss function can calculate and compare the difference between the prediction matrix model and the real matrix model. The smaller the loss function, the better the robustness of the model (Shen et al., 2019). The decline of loss function indicates that the model is in gradient decline to obtain the optimal model.

$$\min_{W\geq 0, H\geq 0} \| X - WH \|^2 . \tag{1}$$

#### Optimization

In the NMF algorithm, we expect to find two matrices W and H. The error between the value of each position corresponding to the matrix obtained by the product of the two matrices and the value of the corresponding position of the original matrix is as small as possible. In the process of NMF model optimization, different loss function models can be better developed. Through NMF model optimization, the generalization ability of the overall model will be greatly improved (Shen et al., 2019). In the formula, W and H represent two different matrices without negative values.

$$\mathrm{W}_{ij} = \mathrm{W}_{ij} \frac{(XH^T)_{ij}}{(WHH^T)_{ij}} . \tag{2}$$

$$H_{ij} = H_{ij} \frac{(W^T X)_{ij}}{(W^T W H)_{ij}} . \tag{3}$$

#### Advantage

NMF is a linear algorithm of geometric structure, which can better reflect the distribution of accurate data. Also, it is a non-parametric algorithm that is very user-friendly. Additionally, it uses a non-parametric method that is very user-friendly. The eigenvalue decomposition issue, which has a globally optimal solution, can also be seen as the counterpart of the NMF algorithm. In the mathematical model, the matrix decomposition image itself is a matrix. It can rely on the mathematical knowledge of matrix decomposition to obtain some special element values and distribution characteristics in this matrix. And these distribution features have good robustness. It can be calculated through this feature and image similarity. The NMF algorithm's primary concept is to divide the non-negative matrix into two distinct matrices, the base matrix and the coefficient matrix, from which the image's essential details can be extracted. In addition, NMF can make a better interpretation of the basis matrix. For example, the NMF method is used to segment human faces. The primary features of the human eyes, nostrils, mouth, and other features are in the obtained basis vector. A weighted combination of these characteristics serves as the source image for the NMF. Thus, the NMF algorithm can be well applied to related face recognition scenes. The NMF algorithm can be effectively used for image decomposition due to the matrix it decomposes has no negative value and negative values have no practical meaning.

### 3.2. $L_{2,1}$-Norm Based NMF Algorithm

Díaz, Steele, and Nguyen indicated that $L_{2,1}$-Norm Based NMF algorithm is an improved algorithm based on the NMF algorithm (Díaz et al., 2021). It is a novel and effective technique that can be used in any area that is susceptible to outliers.

#### Loss Function

The loss function of the $L_{2,1}$-Norm Based NMF algorithm is different from that of the standard NMF algorithm. Since the square residual is removed in this method, the error is not squared in the algorithm, which could increase the robustness against noise.

$$\| X - WH \|^2 = \sum_{i=1}^{n} \| x_i - Wh_i \|^2 . \tag{4}$$

$$\| X - WH \|_{2,1} \Leftarrow \sum_{i=1}^{n} \sqrt{\sum_{j=1}^{p} (X - WH)_{j,i}^2} = \sum_{i=1}^{n} \| x_i - Wh_i \| . \tag{5}$$

#### Optimization

In $L_{2,1}$-Norm Based NMF, the update rules of the NMF need to be adjusted. It calculates $D_{ii}$ in the loss function as the diagonal matrix and then performs data processing (Díaz et al., 2021).

$$\begin{gathered} D_{ii} \Leftarrow \frac{1}{\sqrt{\sum_{j=1}^{p} (X - WH)_{ji}^2}} \\ W_{jk} \Leftarrow W_{jk} \frac{(XDH^T)_{jk}}{(WHDH^T)_{jk}} \\ H_{ki} \Leftarrow H_{ki} \frac{(W^T XD)_{ki}}{(W^T WHD)_{ki}} \end{gathered} \tag{6}$$

### Advantage

$L_{2,1}$-Norm Based NMF algorithm has simple and fast performance. In the practical application of the standard NMF algorithm, a relatively stable objective function is needed when applied to some complex fields (Díaz et al., 2021). However, the $L_{2,1}$-Norm Based NMF can be solved directly with a relatively simple and effective method to obtain the final result.

## 3.3. Noise

### Salt and Pepper Noise

Salt and pepper noise is a kind of impulse noise that destroys part of the original image's pixels (Deng et al., 2016). It uses black pixels with a pixel value of 0 as pepper noise and white pixels with a pixel value of 255 as salt noise to replace some of the pixels in the original image. In the case of a 4*4 matrix, salt and pepper noise converts some pixels in the original matrix to 0 or 255 (Figure 1).

```
[[ 5  6  7  8]
 [ 9 10 11 12]
 [13 14 15 16]
 [17 18 19 20]]
——After Destruction——
[[  5   6   0 255]
 [  0  10 255   0]
 [255 255  15  16]
 [ 17   0  19  20]]
```

**Figure 1.** Principle of salt and pepper noise

This experiment uses randomly produced salt and pepper noise, that the contaminated pixels' location is not fixed. It controls the noise damage level only by the parameter p. For example, when p=0.5, it will destroy 50% of the pixels. The experimental method of adding salt and pepper noise is first to randomly select some pixels and adjust their pixel value to 255. Then randomly select half of the changed pixels and change their pixel value to 0. The resulting contaminated images had a 1:1 ratio of salt noise to pepper noise. Taking an image in the ORL dataset as an example, the following shows the image changes under the damage of different levels of salt and pepper noise (Figure 2).

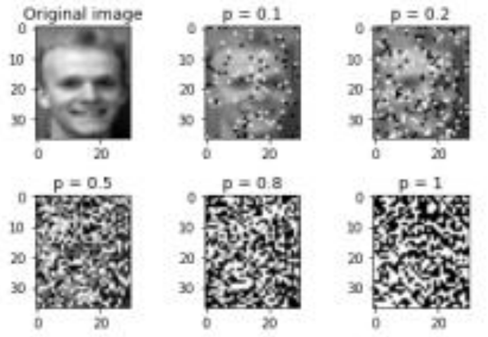


**Figure 2.** Salt and pepper noise damage to images

### Random Matrix Noise

Random matrix noise is named according to its characteristics, which also destroys some pixels in the original image. Unlike salt and pepper noise, this noise does not reset part of the pixels in the original image. It randomly generates a matrix of the same shape as the original image, which is added to the matrix of the original image to form a new contaminated matrix. Take a 4*4 matrix as an example, the pixels in the original matrix are changed in varying degrees (Figure 3).

```
[[ 5  6  7  8]
 [ 9 10 11 12]
 [13 14 15 16]
 [17 18 19 20]]
————————————Noise————————————
[[0.22665339 0.62520188 0.29336913 0.84849946]
 [0.28251995 0.78675625 0.28184821 0.25233337]
 [0.2590281  0.71633493 0.0728687  0.31345217]
 [0.78114883 0.42898609 0.35596487 0.70082194]]
————————After Destruction————————
[[ 5.22665339  6.62520188  7.29336913  8.84849946]
 [ 9.28251995 10.78675625 11.28184821 12.25233337]
 [13.2590281  14.71633493 15.0728687  16.31345217]
 [17.78114883 18.42898609 19.35596487 20.70082194]]
```

**Figure 3.** Principle of random matrix noise

The random matrix noise used in this experiment is also randomly generated, which the damage to each pixel point is not fixed. The parameter p can be used to control the level of noise damage. For example, when p=50, 50 noise matrices will be added to the original image matrix. The method of adding random matrix noise in the experiment is to generate a noise matrix randomly, and each pixel value in the noise matrix is randomly generated between 0 and 1. Then, the original image matrix is added to the noise matrix to obtain a contaminated image matrix. Taking an image in the ORL dataset as an example, the following shows the image changes under the damage of different levels of random matrix noise (Figure 4).

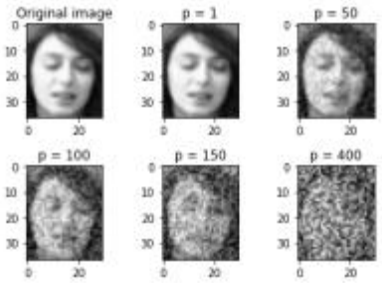


**Figure 4.** Random matrix noise damage to images

## 4. Experiments and Discussions

### 4.1. Datasets

#### ORL Dataset

The ORL dataset is a dataset of 400 images of human faces. The dataset invited different subjects to be photographed at different times, with adjustments made for the light, the subject's facial expressions, and details of their faces (with or without glasses). The size of all images is 92x112 pixels. In the experiment of the NMF algorithm, the image size will bring great computational complexity; so, this experiment reduces the image in ORL dataset by 3 times, scaling the image into a vector containing 1110 elements and the size of the entire image becomes 30*37 pixels.

#### Extended Yale B Dataset

The Extended Yale B Dataset is a dataset of face images containing 2414 images. The dataset invited 38 subjects to photograph in 64 different light conditions with various posture. The size of all images is 168x192 pixels. In order to reduce the pressure of calculation brought by large images, the experiment reduces the image in the Extended Yale B Dataset by 8 times, scaling the image to a vector containing 504 elements and making the entire image become 21*24 pixels.

### 4.2. Algorithm Settings

In order to better ensure the difference between the robustness of the two algorithms, the same experimental environment was set for the two algorithms in this experiment, and only the core operation rules of the two algorithms were kept different. In the initial stage of iterative update, the loss decrease of the two algorithms will occur to a large extent. With the continuous update, the decrease will be lower and lower. To ensure the running time of the whole program, the maximum number of iterations of the algorithm is set as 500 in the experiment. The loss value of the algorithm is displayed every 50 iterations in the experiment. What is more, to ensure that the algorithm does not carry out the meaningless iterative process, the "critical" ESP value is set to 1e-5. When the difference between two iterations is less than the ESP value, the iteration is automatically terminated.

### 4.3. Noise Settings

#### Salt and Pepper Noise

Due to salt and pepper noise directly resets the value of some pixels in the original image to 0 or 225, this noise will cause great damage to the whole image from the visual effect. As the maximum iteration number of 500 is set at a low level in this experiment, to ensure the final effect of the experiment, the basic outline of the picture was retained when selecting the maximum damage level of noise. The p values of salt and pepper noise were set to 0.1, 0.2, and 0.5.

#### Random Matrix Noise

The image damaged by a single pixel in each random matrix noise is between 0 and 1, so a small p value is meaningless for random matrix noise. It is considered that the image's basic contour should be preserved in the experiment; the p values of the random matrix noise were set as 50, 100, and 150.

### 4.4. Evaluation Metrics

#### Relative Reconstruction Errors (RRE)

Relative reconstruction error (RRE) is an excellent way to evaluate NMF algorithms. It is used to represent the similarity between the original data matrix and the reconstructed data matrix. Because of the property of NMF, it will decompose the original large matrix into two small matrices. After the two small matrices are updated, they can be multiplied back to a new matrix with the same size as the original matrix (Díaz et al., 2021). RRE uses this theory to compare errors between two sets of data.

$$RRE = \frac{\| \hat{X} - WH \|}{\| \hat{X} \|}. \tag{7}$$

In the above formula, $\hat{x}$ represents clean data, and U and V represent the decomposition results on $\hat{x}$ respectively.

#### Accuracy

In the prediction and reconstruction of the data matrix, each image data set contains different subjects. We will calculate the accuracy of clustering experiment and obtain the average accuracy and standard deviation of data. Average accuracy and standard deviation can reflect the variation trend and difference between the prediction and original data models.

#### Relative Reconstruction Errors (RRE)

Assuming that the standard clustering result is 'x' and 'o' in the figure, while the result of our clustering is the large circle, NMI is used to measure the similarity between these results to judge the difference between different data (Figure 5).

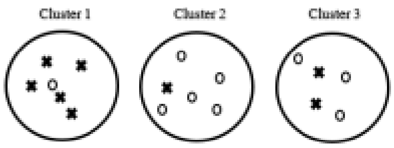


**Figure 5.** Example of NMI Measurement

(A, B) in the formula is the mutual information of two vectors A and B, and H(A) and H(B) are the information entropy of A vector and B vector.

$$NMI = \frac{H(A) + H(B)}{H(A,B)} . \quad (8)$$

## 4.5. Results and Analysis

The experiment's primary purpose is to compare the robustness differences between the standard NMF algorithm and the $L_{2,1}$-Norm Based NMF under different noise disturbances. Due to the complexity of the calculation caused by the insertion of a large amount of random noise into the image, the final clustering result is often haphazard. In order to ensure the reliability of experimental results, this experiment randomly extracts 90% images from the dataset for the clustering test five consecutive times and evaluate the two algorithms using the mean and standard deviation of the final results.

### Salt and Pepper Noise Experiment

The experiment was first tested using salt and pepper noise on ORL and Yale B datasets. By adjusting the proportion of noise pixels in the image, we obtained the tendency of the mean values of the three evaluated metrics set by the experiment to change with the noise proportion (Figure 6 & Figure 7).

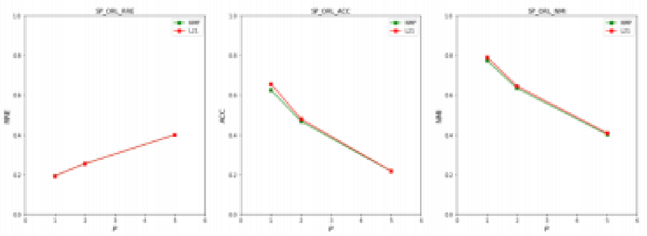

**Figure 6.** Experiment using salt and pepper noise on ORL dataset

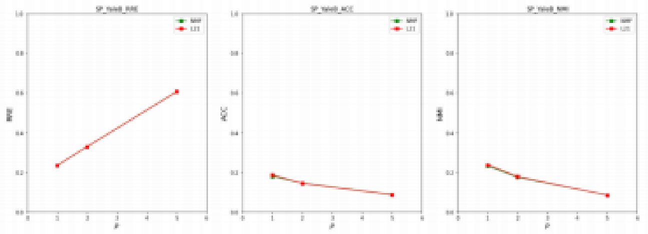

**Figure 7.** Experiments using salt and pepper noise on the Yale B dataset

The experimental findings show that the RRE values of the two algorithms in the various datasets exhibit a rising trend as the p value increases, while the accuracy and NMI value significantly decline. It is because with the increase of noise, more pixels in the image are changed, making the algorithm more challenging to reconstruct the image, the clustering effect is decreased. As an example of ORL data set reconstruction using the standard NMF algorithm, the following figure shows the differences in reconstructed images under different salt and pepper noise levels (Figure 8 & Figure 9).

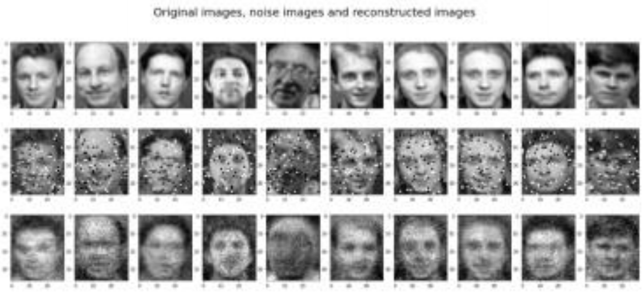


**Figure 8.** Influence of salt and pepper noise on reconstructed image when p=0

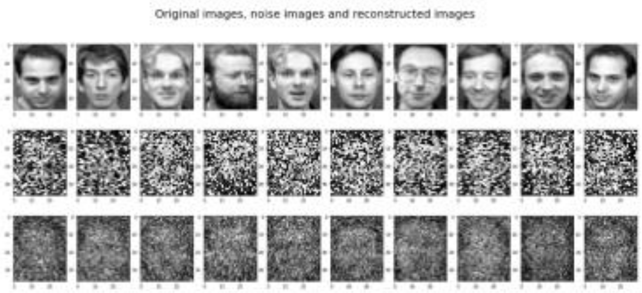


**Figure 9.** Influence of salt and pepper noise on reconstructed image when p=0.5

Judging from the visual effect, when p=0.1, the contaminated image still retains some original image details. Since some details in the original picture can more easily be restored, the reconstructed image resembles the original image more. However, when p=0.5, most of the original image details are destroyed, making it more difficult for the two NMF algorithms to restore more details in the image reconstruction. It will increase the difficulty of clustering, resulting in the final clustering accuracy and similarity decreased.

The line graph shows that the two NMF algorithms produce relatively similar performance under the interference of salt and pepper noise, especially in the graph related to the Yale B dataset, the line representing the two NMF algorithms almost overlaps. Therefore, the mean and standard deviation of test results were recorded in tables for further exploration (Table 1 & Table 2).

Table 1. Average results of salt and pepper noise experiment

| **Mean** | | **ORL** | | | **YaleB** | |
|---|---|---|---|---|---|---|
| **Salt Pepper, p = 0.1** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | 0.1931 | 0.6256 | 0.7738 | 0.2331 | 0.1794 | 0.2312 |
| **$L_{2,1}$-Norm Based NMF** | **0.1923** | **0.6572** | **0.7915** | **0.2330** | **0.1854** | **0.2373** |
| **Salt Pepper, p = 0.2** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | **0.2535** | 0.4683 | 0.6348 | 0.3278 | **0.1448** | 0.1746 |
| **$L_{2,1}$-Norm Based NMF** | 0.2538 | **0.4794** | **0.6453** | **0.3276** | 0.1437 | **0.1774** |
| **Salt Pepper, p = 0.5** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | 0.3996 | 0.2167 | 0.4019 | 0.6055 | **0.0868** | **0.0848** |
| **$L_{2,1}$-Norm Based NMF** | **0.3993** | **0.2178** | **0.4087** | **0.6053** | 0.0863 | 0.0842 |

Table 2. Standard deviation of salt and pepper noise experiment results

| **Standard Deviation** | | **ORL** | | | **YaleB** | |
|---|---|---|---|---|---|---|
| **Salt Pepper, p = 0.1** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | **0.0005** | 0.0203 | 0.0163 | 0.0009 | 0.0105 | 0.0073 |
| **$L_{2,1}$-Norm Based NMF** | 0.0006 | **0.0202** | **0.0120** | **0.0008** | **0.0080** | **0.0065** |
| **Salt Pepper, p = 0.2** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | **0.0005** | 0.0263 | 0.0197 | 0.0017 | 0.0056 | **0.0028** |
| **$L_{2,1}$-Norm Based NMF** | 0.0007 | **0.0163** | **0.0086** | **0.0011** | **0.0025** | 0.0049 |
| **Salt Pepper, p = 0.5** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | 0.0014 | **0.0093** | 0.0158 | 0.0028 | **0.0021** | **0.0035** |
| **$L_{2,1}$-Norm Based NMF** | **0.0009** | 0.0129 | **0.0145** | **0.0015** | 0.0028 | 0.0044 |

The table showing the average value of the test findings demonstrates that as the degree of salt and pepper noise damage increases, the performance of $L_{2,1}$-Norm Based NMF algorithm on ORL dataset is better than that of standard NMF algorithm. This algorithm has a lower RRE value, higher clustering accuracy and clustering similarity. Meanwhile, by referring to the table representing the standard deviation of test results, we find that the $L_{2,1}$-Norm Based NMF algorithm is stable in most cases for ORL dataset.

In the Yale B dataset, the performance of the two algorithms has significantly decreased. It may be because the Yale B dataset itself is too complex, and a large scale is used when scaling images, making it difficult to perform accurate clustering on the data set itself. In the case of relatively low noise damage, the $L_{2,1}$-Norm Based NMF algorithm still maintains better performance and strong stability. When the degree of noise damage increases, the standard NMF algorithm is superior to $L_{2,1}$-Norm Based NMF algorithm in clustering accuracy and similarity, but the difference between the two algorithms is minimal. Considering that the clustering effect of this dataset is generally low, these minimal differences are considered to have no tremendous reference value under the interference of excessive noise.

Finally, in the experiments to evaluate the robustness of the two algorithms in the face of salt and pepper noise, we learned that salt and pepper noise significantly interfered with the NMF algorithm, and the performance of the two algorithms declined significantly under the interference of high-intensity salt and pepper noise. Compared with the two algorithms, $L_{2,1}$-Norm Based NMF algorithm has better overall performance than the standard NMF algorithm in terms of accuracy and stability.

## Random Matrix Noise Experiment

Random matrix noise was also tested on ORL and Yale B datasets. By adjusting the multiples of the noise pixels applied to the image, we obtained the tendency of the mean values of the three evaluated metrics set by the experiment to change with the noise ratio (Figure 10 & Figure 11).

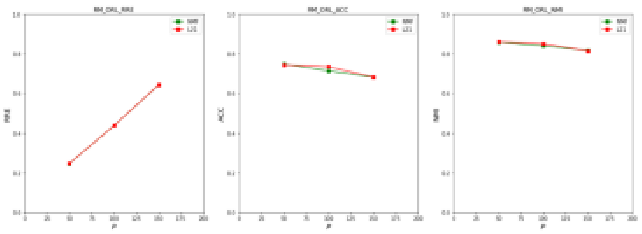

**Figure 10.** Experiment using random matrix noise on ORL dataset

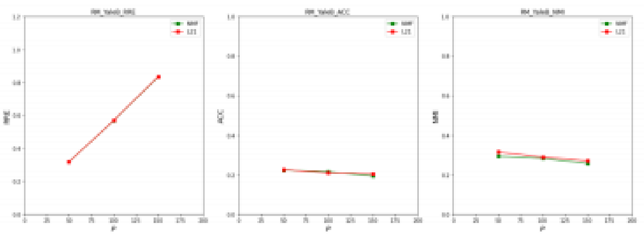

**Figure 11.** Experiments using random matrix noise on the Yale B dataset

The experimental results show that with the increase of p value, the RRE value of the two algorithms in different datasets presents a rising trend, while the accuracy and NMI value begins to decline. Compared with salt and pepper noise, random matrix noise resulted in a more significant increase in RRE values and a smaller decrease in accuracy and NMI values. Also, taking ORL dataset reconstruction using standard NMF algorithm as an example, the following figure shows the differences of reconstructed images under different levels of random matrix noise (Figure 12 & Figure 13).

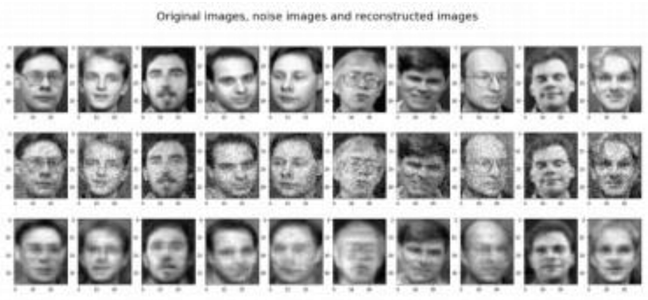


**Figure 12.** Influence of random matrix noise on reconstructed image when p=50

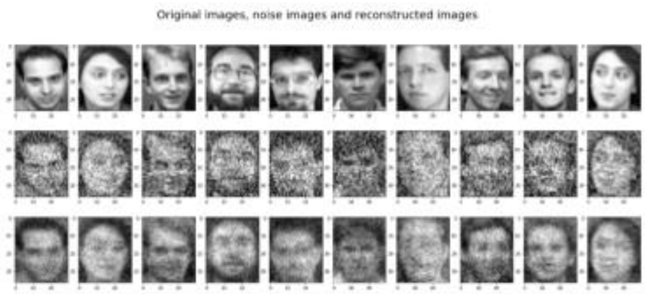


**Figure 13.** Influence of random matrix noise on reconstructed image when p=150

Judging from the visual effect, when the p value changes from 50 to 150, the definition of the contaminated picture is significantly reduced, but it still retains many details in the original picture. It is easier for the NMF algorithm to restore most dataset features during image reconstruction. Although the reconstructed image is still unclear, it enables the two NMF algorithms to achieve higher and more stable clustering accuracy and similarity on random matrix noise. The significant increase in RRE value is that the fuzzy matrix algorithm modifies all pixels in the original image, which leads to specific changes in almost all pixels compared with the original image in the reconstruction of the image. The overall change does not have a significant visual impact on the reconstruction of local details, but it will lead to more significant differences in the overall values of the two image matrices.

Similar to the salt and pepper noise experiment, the standard NMF algorithm and $L_{2,1}$-Norm Based NMF algorithm produce similar performance under the interference of random matrix noise. Especially when testing the Yale B dataset, it is not easy to see and get the difference between the two algorithms by looking at the line graph. Therefore, the mean and standard deviation of test results were recorded in tables for further exploration. Therefore, the mean and standard deviation of test results were recorded in tables for further exploration (Table 3 & Table 4)

.

Table 3. Average value of random matrix noise experiment results

| Mean | ORL | | | YaleB | | |
|---|---|---|---|---|---|---|
| **Random Matrix, p = 0.1** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | **0.2451** | **0.7478** | 0.8595 | 0.3167 | 0.2233 | 0.2928 |
| **$L_{2,1}$ -Norm Based NMF** | 0.2454 | 0.7439 | **0.8605** | **0.3166** | **0.2250** | **0.3142** |
| **Random Matrix, p = 0.2** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | 0.4378 | 0.7144 | 0.8407 | **0.5685** | **0.2141** | 0.2822 |
| **$L_{2,1}$ -Norm Based NMF** | **0.4377** | **0.7350** | **0.8512** | 0.5697 | 0.2094 | **0.2886** |
| **Random Matrix, p = 0.5** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | **0.6426** | 0.6822 | **0.8178** | **0.8341** | 0.1938 | 0.2561 |
| **$L_{2,1}$ -Norm Based NMF** | 0.6434 | **0.6828** | 0.8150 | 0.8365 | **0.2043** | **0.2708** |

Table 4. Standard deviation of random matrix noise experiment results

| Standard Deviation | ORL | | | YaleB | | |
|---|---|---|---|---|---|---|
| **Random Matrix, p = 0.1** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | **0.0005** | 0.0194 | 0.0068 | 0.0012 | 0.0078 | **0.0170** |
| **$L_{2,1}$ -Norm Based NMF** | **0.0005** | **0.0159** | **0.0062** | **0.0008** | **0.0074** | 0.0195 |
| **Random Matrix, p = 0.2** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | **0.0006** | **0.0241** | 0.0155 | 0.0009 | 0.0105 | **0.0122** |
| **$L_{2,1}$ -Norm Based NMF** | **0.0006** | 0.0287 | **0.0149** | **0.0010** | **0.0040** | 0.0092 |
| **Random Matrix, p = 0.5** | **RRE** | **ACC** | **NMI** | **RRE** | **ACC** | **NMI** |
| **Standard NMF** | 0.0012 | 0.0231 | 0.0179 | **0.0015** | 0.0088 | **0.0146** |
| **$L_{2,1}$ -Norm Based NMF** | **0.0009** | **0.0159** | **0.0154** | 0.0019 | **0.0069** | 0.0158 |

Both NMF algorithms outperformed the salt and pepper noise experiment on two datasets that contained random matrix noise. When p=50, the standard NMF algorithm shows better accuracy in the ORL dataset, but the corresponding stability is lower than the $L_{2,1}$-Norm Based NMF algorithm. In the Yale B dataset, when p=100, the standard NMF algorithm shows better accuracy, but the corresponding stability is still lower than $L_{2,1}$-Norm Based NMF algorithms. In most other cases, $L_{2,1}$-Norm Based NMF algorithms outperform standard NMF algorithms on both datasets.

In the experiments to evaluate the robustness of the two algorithms in the face of random matrix noise, we learned that the interference of random matrix noise to the NMF algorithm is minor, and the performance of the two algorithms has been improved compared with that of salt and pepper noise. Compared with the two algorithms, $L_{2,1}$-Norm Based NMF algorithm has better overall performance than the standard NMF algorithm in terms of accuracy and stability.

## 5. Conclusions and Future Work

### 5.1. Conclusions

In this study, we chose the standard NMF algorithm and the $L_{2,1}$-Norm Based NMF algorithm as the research object. By adding salt and pepper noise and random matrix noise to the ORL dataset and the Extended Yale B dataset, we designed the clustering experiment and compared the robustness of the two NMF algorithms.

According to the experimental results, it can be concluded that compared with salt and pepper noise, the random matrix noise has a minor impact on the clustering test of the two NMF algorithms. However, since the random matrix noise will modify all the pixels, it will cause a more significant relative reconstruction error. For the same kind of noise, more serious noise will significantly reduce the efficiency of NMF algorithm in reconstructing the image to restore details. As a result, the accuracy and similarity of clustering will decrease with the increase of noise severity.

When comparing the two NMF algorithms, $L_{2,1}$-Norm Based NMF perform better than NMF in most cases in the face of different noise and datasets. $L_{2,1}$-Norm Based NMF algorithms generally provide more accurate and stable results. However, in this experiment, the maximum difference between the three evaluation indexes of the two algorithms under the same conditions is within 0.05. In conclusion, in this experiment, the $L_{2,1}$-Norm Based NMF algorithm and standard NMF algorithm show very similar robustness, but the $L_{2,1}$-Norm Based NMF algorithm has a slight advantage.

## 5.2. Future Work

In this study, only the standard NMF algorithm and $L_{2,1}$-Norm Based NMF algorithm were explored. Certain limitations were imposed on the processing of datasets and the setting of the experimental environment, so the experimental results had certain restrictions. In the future research, we will conduct further research from the following aspects:

- Choose more datasets in the experiment and retain more original information of the image through a lower zoom ratio, exploring whether a complete image can effectively improve the efficiency of the NMF algorithm.
- Increase the maximum number of iterations in the experiment, exploring whether NMF algorithm can better improve the clustering effect after more multiplication updates.
- Use more NMF algorithms, exploring the difference in robustness between different NMF algorithms.

## References


[1] Deng, X., Ma, Y., & Dong, M. (2016). A new adaptive filtering method for removing salt and pepper noise based on multilayered PCNN. *Pattern Recognition Letters*, *79*, 8-17.

[2] Díaz, A., & Steele, D. (2021). Analysis of the robustness of NMF algorithms. *arXiv preprint arXiv:2106.02213*.

[3] Kong, D., Ding, C., & Huang, H. (2011, October). Robust nonnegative matrix factorization using l21-norm. In *Proceedings of the 20th ACM international conference on Information and knowledge management* (pp. 673-682).

[4] Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature*, *401*(6755), 788-791.

[5] Shen, X., Zhang, X., Lan, L., Liao, Q., & Luo, Z. (2019). Another robust NMF: rethinking the hyperbolic tangent function and locality constraint. *IEEE Access*, *7*, 31089-31102.