Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14145
e- ISSN : 2541-2019
p- ISSN : 2541-044X

Complete Kernel Fisher Discriminant (CKFD) and Color Difference Histogram for Palm Disease

Johanes Terang Kita Perangin Angin[1)*], Herman[2)], Joni[3)]
[1)2)3)]Information Study Program, STMIK TIME, Medan, Indonesia
[1)]timejohanes@gmail.com, [2)]cubitsakit@gmail.com, [3)]joni_hgw@yahoo.com

Submitted :Sep 29, 2024 | **Accepted** : Oct 31, 2024 | **Published** : Nov 2, 2024

Abstract: Palm oil plantations play a significant role in the economy of Indonesia, supporting 16.2 million people. However, plant diseases pose a major threat to the productivity and health of palm oil crops. Early detection of these diseases is essential to prevent yield losses and mitigate damage. This study proposes the application of the Complete Kernel Fisher Discriminant (CKFD) method combined with Color Difference Histogram to classify diseases affecting oil palm fronds and leaves. The CKFD method uses a non-linear kernel transformation to improve the performance of Fisher Linear Discriminant Analysis (FLDA), while the Color Difference Histogram enhances sensitivity to color variations in different lighting conditions. Experimental results demonstrate that the CKFD method achieves superior accuracy in disease detection compared to traditional Convolutional Neural Networks (CNN) and Support Vector Machines (SVM). The proposed approach showed an average accuracy of 94.5% for detecting diseases like Curvularia sp and Cochliobolus carbonus. The combination of CKFD with Color Difference Histogram significantly reduces the impact of lighting variations on the classification results, making it a robust solution for practical deployment in palm oil plantations. This research provides an effective tool for early disease detection and management in the palm oil industry.

Keywords: Color Difference Histogram; Complete Kernel Fisher Discriminant; Computer Vision; Histogram Neighboring Edge; Histogram Neighboring Color Indexes

INTRODUCTION

The development of the palm oil plantation sector is an integral part of national development. Based on the Decree of the Minister of Agriculture Number 833/KPTS/SR. 020/M/12/2019, the area of palm oil plantations in Indonesia reaches 16.38 million hectares spread across 25 provinces. Palm oil plantations are the backbone of the economy for 16.2 million Indonesians(Satia et al., 2022). Frandian explains that pest or disease attacks on the fronds and leaves of oil palm plants will inhibit the growth process of oil palms(Frandian et al., 2022), so that plant growth becomes stunted, there is a decrease in productivity, and even the plants can die. Hamdani et al. has applied the principles of computer vision in analyzing color histograms of oil palm fronds and leaves images to detect oil palm plant diseases(Hamdani et al., 2021). Sarkar et al. have applied computer vision with machine learning in analyzing diseases in oil palm plants based on leaf images(Sarkar et al., 2017). The same research has been conducted by a number of researchers who analyzed images of oil palm leaves using CNN(Kipli et al., 2023; Matarneh et al., 2024; Panchal et al., 2023). One of the accurate computer vision methods for detecting images based on images of leaf stalks and oil palm plants is based on color histograms(Hamdani et al., 2021; Hasan et al., 2022). The application of computer vision to overcome color histograms has a major problem in the form of complexity in the existing color features(Singh et al., 2018). A fairly good method for analyzing color histograms in an image is the Local Fisher Discriminant Analysis (LFDA) method(Jia et al., 2016). LFDA is considered to be the method that provides the best results in image identification(Wang et al., 2016). The LFDA method has a major problem in the form of low accuracy on complex features, so the Complete Kernel Fisher Discriminant (CKFD) method was proposed(Dong et al., 2016; Wen et al., 2018; Zafeiriou, 2012). The CKFD method based on the results of previous research conducted by researchers has provided good accuracy in facial image recognition but is sensitive to changes in color histograms caused by lighting(Angin et al., 2020). The application of Color Difference Histogram which adds Histogram Neighboring Edge and Histogram Neighboring Color Indexes so that

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14145
e- ISSN : 2541-2019
p- ISSN : 2541-044X

it can provide accuracy in the changes in color histogram in test data which may be different from the training data(Liu & Yang, 2013).

LITERATURE REVIEW

CKFD Method

The Complete Kernel Fisher Discriminant Analysis (CKFD) method applies a non-linear kernel trick before applying the FLDA discriminant rule. The CKFD method was originally intended to overcome the complexity that occurs in facial images so that it still leaves a singularity problem (small training sample problem). Efforts to improve CKFD performance can be made in various ways, namely by kernel innovation, new discriminant rules, the use of different classifiers, or a combination of them. In this study, the use of the Spectrally Weighted (SW) kernel derived from the Polynomial Kernel and Radial Based Function (RBF) in the SVM classifier mechanism was studied. By applying the LibSVM decomposition algorithm based on gradient descent, it is expected to obtain weight optimization that can increase the performance of the recognition rate(Dong et al., 2016; Wen et al., 2018; Zafeiriou, 2012).

The Complete Kernel Fisher Discriminant (CKFD) method has two advantages when compared to the previous Kernel Fisher Discriminant (KFD). First, the implementation of this algorithm can be divided into two phases, namely Kernel principal component analysis (KPCA) plus Fisher linear discriminant analysis (FLD) so that the results are more transparent and simpler. Second, CKFD can create two categories of discriminant information so that the results are stronger.

In the KPCA space transformation, it can be divided into two classes, namely between class and within class scatter matrices(S_b, S_w).The equations used can be seen in Equations 1 and 2.

$$S_b = \sum_{i=1}^{e} P(\omega_i)(m_i - m_o)(m_i - m_o)^T \tag{1}$$

$$S_w = \sum_{i=1}^{e} P(\omega_i)\left\{\sum_{j=1}^{l_i} \frac{1}{l_i}(X_{ij} - m_i)(X_{ij} - m_i)^T\right\} \tag{2}$$

Where:

X_{ij} : Denote *j* as a training sample of class *i*.
P(ωi) : Probability of Class *i*.
Li : The number of training samples in class *i*
M_i : The average of the training sample vectors in class *i*.
M_o : The mean vector is the mean vector of all training samples.

Color Difference Histogram Method

The application of Color Difference Histogram adds Histogram Neighboring Edge and Histogram Neighboring Color Indexes so that it can provide accuracy in the color histogram changes in the test data which may be different from the training data.

The Neighboring Edge Histogram Vector can be determined using Equation 3.

$$H_{Color}\big(Q(X_c,Y_c)\big) = \sum\sum\sqrt{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2}, \theta(X_c, Y_c = \theta(X,Y) \tag{3}$$

Vector Histogram Neighboring Color Indexes can be determined using Equation 4.

$$H_{Orient}\big(Q(X_c,Y_c)\big) = \sum\sum\sqrt{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2}, Q(X_c, Y_c = Q(X,Y) \tag{4}$$

Where the two vectors are combined using Equation 5.

$$H_{CDH} = \{H_{color}(0), .., H_{color}(n), H_{Orient}(0), \ldots, H_{Orient}(n)\} \tag{5}$$

Palm Oil Disease

Palm Oil disease can be seen in Table 1.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024 e- ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v8i4.14145 p- ISSN : 2541-044X

Table 1. Palm Oil Disease

No	Picture	Name	Characteristic
1		Curvularia sp	This disease is characterized by the appearance of small, round, yellow spots on the leaves. These spots can be seen on both sides of the leaf surface and are surrounded by a yellowish-orange layer.
2		Cochliobolus carbonus	This disease causes brown, bird-eye-shaped spots. The diameter of the spots is small, about ≤0.5 mm, and the halo (circle around the spot) is 5-7 mm. Sometimes, these spots can infect the entire leaf.
3		Capnodium sp	Leaves affected by this disease will be coated with a black layer, making it look like soot.
4		Drecshlera	The resulting spots are yellowish brown to blackish in color on the leaves of the plant.
5		Nutrient Defficiency	Symptoms of this deficiency include leaves at the tip of the leaf stalk curling and experiencing chlorosis (yellowing of the leaves) followed by necrosis (death of leaf tissue) at the tip of the leaf. If the condition is severe, the entire leaf will experience necrosis and eventually hang down.

METHOD

Research Method

The research method can be seen in Figure 1.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14145
e- ISSN : 2541-2019
p- ISSN : 2541-044X

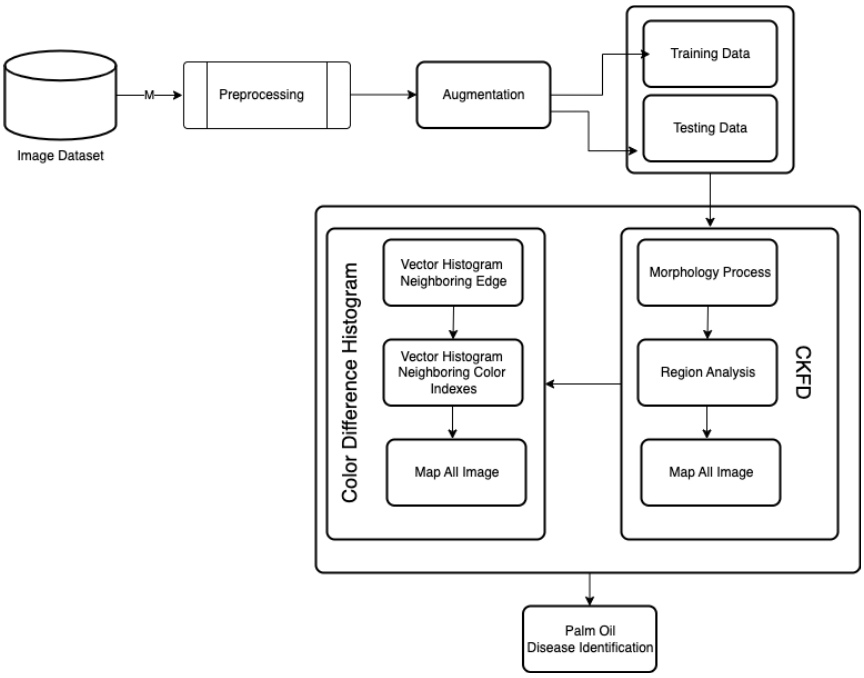

Fig. 1. Research Method

Based on Figure 1, it can be seen that The Complete Kernel Fisher Discriminant (CKFD) method with Color Difference Histogram comprises the following stages:

Preprocessing

In this stage, input images are preprocessed to enhance their quality and standardize their format for further analysis. This often includes resizing, noise reduction, and normalization techniques. The purpose of preprocessing is to ensure that the input data is consistent and that any irrelevant noise is minimized, improving the efficiency of feature extraction.

Augmentation

Data augmentation techniques are applied to increase the size and diversity of the dataset. This involves operations like rotation, scaling, cropping, and brightness adjustment. The augmented data helps the CKFD model generalize better by training on various image conditions, thereby reducing overfitting.

Determination of Training and Testing Data

The dataset is divided into training and testing subsets, usually in a ratio of 70:30 or 80:20. The training set is used to train the CKFD model, while the testing set is reserved for evaluating the model's performance. This step is crucial for ensuring that the model is tested on unseen data and can provide a realistic estimate of its generalization ability.

CKFD Method

The CKFD method involves the following processes:

a. **Morphology Process**: Morphological operations like dilation, erosion, opening, and closing are used to enhance the structures within the image. These operations help in refining the shapes and boundaries of objects, making it easier to extract meaningful features.
b. **Region Analysis**: The image is divided into different regions based on intensity or color. These regions are analyzed individually to identify key features that distinguish one region from another. Region analysis helps in capturing local features that may not be evident at the global image level.
c. **Map All Image**: The entire image is mapped based on the extracted regional features. This mapping process creates a comprehensive feature representation of the image, which is then passed to the Color Difference Histogram stage for further processing.

Color Difference Histogram

The Color Difference Histogram is used to capture color-based features from the image. It consists of the following sub-stages:

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14145
e- ISSN : 2541-2019
p- ISSN : 2541-044X

a. **Vector Histogram Neighboring Edge**: This sub-stage focuses on calculating the histogram of edges based on neighboring pixels. The histogram captures the distribution of edge intensities and directions, which provides crucial structural information about the image.
b. **Vector Histogram Neighboring Color Indexes**: Histograms are created based on the neighboring color indexes, capturing the color distribution and differences between adjacent pixels. This step provides a detailed representation of the color composition within the image.
c. **Map All Image**: The image is mapped using the histograms generated from the previous sub-stages. This final mapping combines color and structural information, creating a comprehensive feature map that is used for classification or further analysis.

Proposed Method

Proposed method can be seen in the pseudocode as follows.

1. Input: Image Dataset
2. Preprocessing:
   - For every pixel (*x*,*y*) in Image *I:*
     - Convert the RGB values to L*a*b* color space
     - Compute the color difference between pixel (x, y) and its neighboring pixels using Equation 6.

$$color\_diff = \sqrt{(DL)^2 + (Da)^2 + (Db)^2} \quad (6)$$

     - Store the color difference in Histogram
1. Processing:
   - For each pixel (*x*,*y*) in Image *I*:
     - Compute the gradient for each channel (L*, a*, b*) using the Sobel operator
     - Compute the edge orientation θ using Equation 7.

$$\theta = 0.5 * arctan\left(\frac{2.g_{xy}}{g_{xx}-g_{yy}}\right) \quad (7)$$

     - Store the edge orientation in the histogram
   - For each image *I* in training dataset
     - Compute the Scatter matrix (S_b and S_w) using equation 8 and 9.

$$S_b = \sum(m_i - m_o)(m_i - m_o)^T \quad (8)$$

$$S_w = \sum\sum(x_{ij} - m_i)(x_{ij} - m_i)^T \quad (9)$$

     - Apply Kernel Principal Component Analysis (KPCA) transformation.
     - Apply Fisher Linear Discriminant Analysis (FLDA) for feature discrimination.
     - Compute the Eigen Vectors and sort them based on the highest eigenvalues.
   - For each image *I* in testing dataset
     - Compute the Euclidean distance between the test image and the training images using Equation 10.

$$distance = \sqrt{\sum(u[i] - t[i])^2} \quad (10)$$

     - Identify the face based on the smallest distance.

The pseudocode provided outlines a step-by-step process for feature extraction and classification of images, particularly focusing on color and gradient features. The first step involves converting the input image's RGB values to the Lab* color space for better color representation. For each pixel(*x*,*y*) in the image, the algorithm computes the color difference between that pixel and its neighboring pixels, which is then stored in a histogram. In the processing stage, the gradient for each channel (L*, a*, and b*) is computed using the Sobel operator, and the edge orientationθ is calculated. These values are stored in a histogram as well. For the training images, scatter matrices (Sb and Sw) are calculated to represent between-class and within-class scatter using predefined equations. The algorithm then applies Kernel Principal Component Analysis (KPCA) for dimensionality reduction and Fisher Linear Discriminant Analysis (FLDA) for feature discrimination. Eigenvectors are computed and sorted based on the highest eigenvalues. For each test image, the Euclidean distance between the test image and the training images is computed, and the image is classified based on the smallest distance. This approach effectively leverages color and gradient information for image classification.

RESULT

This study will compare the results obtained by the CKFD method with Color Difference Histogram compared to the CNN and SVM methods. The results obtained can be seen in Table 2.

Table 2. Results

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14145
e- ISSN : 2541-2019
p- ISSN : 2541-044X

Leaf Disease	Method	Accuration (%)	Precision (%)	Recall (%)
Curvularia sp	Complete Kernel Fisher Discriminant (CKFD)	94.5	93.2	95.1
Curvularia sp	Convolutional Neural Network (CNN)	90.2	89.0	91.8
Curvularia sp	Support Vector Machine (SVM)	88.6	87.0	90.1
Cochliobolus carbonus	Complete Kernel Fisher Discriminant (CKFD)	92.7	91.5	94.0
Cochliobolus carbonus	Convolutional Neural Network (CNN)	87.5	85.8	89.2
Cochliobolus carbonus	Support Vector Machine (SVM)	85.9	84.2	87.7
Embun Jelaga (Capnodium sp)	Complete Kernel Fisher Discriminant (CKFD)	95.3	94.0	96.5
Embun Jelaga (Capnodium sp)	Convolutional Neural Network (CNN)	92.1	90.5	93.8
Embun Jelaga (Capnodium sp)	Support Vector Machine (SVM)	89.1	87.7	90.5
Drechslera	Complete Kernel Fisher Discriminant (CKFD)	93.8	92.6	95.0
Drechslera	Convolutional Neural Network (CNN)	89.4	88.2	90.7
Drechslera	Support Vector Machine (SVM)	87.8	86.4	89.3
Defisiensi Unsur Hara	Complete Kernel Fisher Discriminant (CKFD)	91.6	90.1	93.0
Defisiensi Unsur Hara	Convolutional Neural Network (CNN)	88.3	86.9	89.7
Defisiensi Unsur Hara	Support Vector Machine (SVM)	86.2	85.0	87.4

Based on the analysis results shown in table 2, the Complete Kernel Fisher Discriminant (CKFD) method proved to be superior in detecting oil palm leaf diseases compared to other methods, such as CNN and SVM. CKFD showed consistently high accuracy for all types of leaf diseases, especially for Sooty Mildew and Curvularia sp, with accuracies reaching 95.3% and 94.5%. In addition, the precision and recall of CKFD were also very high, respectively indicating that this method was able to accurately classify diseases and detect most cases with a very low error rate. In comparison, the Convolutional Neural Network (CNN) and Support Vector Machine (SVM) methods performed slightly lower than CKFD. CNN, which is usually relied on for image classification, in some cases showed lower accuracy and recall, especially for diseases such as Cochliobolus carbonus and Nutrient Deficiency. The SVM method also had the lowest performance among the three, with lower average accuracy and difficulty in handling the complexity of leaf images. This shows that CKFD is a more effective option in detecting various types of oil palm leaf diseases.

DISCUSSIONS

Based on the results obtained, it is evident that the Complete Kernel Fisher Discriminant (CKFD) method outperforms the Convolutional Neural Network (CNN) and Support Vector Machine (SVM) methods in detecting oil palm leaf diseases. The CKFD method consistently achieved higher accuracy, precision, and recall across various disease types such as Curvularia sp, Cochliobolus carbonus, and Sooty Mildew. The ability of CKFD to handle the complexity of color histograms and maintain high performance even under varying lighting conditions demonstrates its robustness and effectiveness for image classification tasks. The addition of the Color Difference Histogram further enhances its performance by providing more discriminative features for differentiating healthy and diseased leaves.

Compared to traditional CNN and SVM methods, which showed limitations in dealing with the variability and complex features in leaf images, CKFD proved to be a more reliable option. The CNN method, while generally strong in image classification, struggled with certain diseases such as Cochliobolus carbonus and Nutrient Deficiency, where its accuracy and recall were lower. Similarly, SVM, which is typically used for linear classification tasks, was less effective in capturing the intricate patterns in the dataset. These findings suggest that CKFD, combined with the Color Difference Histogram, offers a superior solution for early detection and classification of oil palm leaf diseases, making it a valuable tool for improving disease management in palm oil plantations.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14145
e- ISSN : 2541-2019
p- ISSN : 2541-044X

CONCLUSION

The research likely concludes that the Complete Kernel Fisher Discriminant (CKFD) method is effective for dimensionality reduction and feature extraction, enhancing the separability of features related to different palm diseases. The study may demonstrate that the Color Difference Histogram provides an effective way to capture and quantify color variations in palm images, which is critical for identifying disease-specific patterns.Combining CKFD with color histograms likely enhances classification accuracy for detecting palm diseases, making it a reliable method for early diagnosis.The findings could support applications in precision agriculture, where automated disease detection is key for efficient crop management and intervention.

ACKNOWLEDGMENT

This research is funded by the Ministry of Education, Culture, Research, and Technology of the Republic of Indonesia.

REFERENCES

Angin, J. T. K. P., Johan, Sukiman, Sugianto, Simarmata, B. R., & Suharjito. (2020). Face Recognition Application with the Complete Kernel Fisher Discriminant (CKFD) Method. *2020 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT)*, 359–364. https://doi.org/10.1109/MECnIT48290.2020.9166682

Dong, S., Wang, Z., & Zeng, L. (2016). Lithology identification using kernel Fisher discriminant analysis with well logs. *Journal of Petroleum Science and Engineering*, *143*, 95–102. https://doi.org/10.1016/j.petrol.2016.02.017

Frandian, B., Zufria, I., & Irawan, M. D. (2022). Implementasi Certainty Factor Untuk Diagnosis Penyakit dan Hama Pada Pelepah dan Daun Kelapa Sawit Beserta Penanganannya. *Journal of Information System Research (JOSH)*, *3*(3), 159–168. https://doi.org/10.47065/josh.v3i3.1455

Hamdani, H., Septiarini, A., Sunyoto, A., Suyanto, S., & Utaminingrum, F. (2021). Detection of oil palm leaf disease based on color histogram and supervised classifier. *Optik*, *245*, 167753. https://doi.org/10.1016/j.ijleo.2021.167753

Hasan, S., Jahan, S., & Islam, Md. I. (2022). Disease detection of apple leaf with combination of color segmentation and modified DWT. *Journal of King Saud University - Computer and Information Sciences*, *34*(9), 7212–7224. https://doi.org/10.1016/j.jksuci.2022.07.004

Jia, J., Ruan, Q., & Jin, Y. (2016). Geometric Preserving Local Fisher Discriminant Analysis for person re-identification. *Neurocomputing*, *205*, 92–105. https://doi.org/10.1016/j.neucom.2016.05.003

Kipli, K., Osman, S., Joseph, A., Zen, H., Awang Salleh, D. N. S. D., Lit, A., & Chin, K. L. (2023). Deep learning applications for oil palm tree detection and counting. *Smart Agricultural Technology*, *5*, 100241. https://doi.org/10.1016/j.atech.2023.100241

Liu, G.-H., & Yang, J.-Y. (2013). Content-based image retrieval using color difference histogram. *Pattern Recognition*, *46*(1), 188–198. https://doi.org/10.1016/j.patcog.2012.06.001

Matarneh, S., Elghaish, F., Pour Rahimian, F., Abdellatef, E., & Abrishami, S. (2024). Evaluation and optimisation of pre-trained CNN models for asphalt pavement crack detection and classification. *Automation in Construction*, *160*, 105297. https://doi.org/10.1016/j.autcon.2024.105297

Panchal, A. V., Patel, S. C., Bagyalakshmi, K., Kumar, P., Khan, I. R., & Soni, M. (2023). Image-based Plant Diseases Detection using Deep Learning. *Materials Today: Proceedings*, *80*, 3500–3506. https://doi.org/10.1016/j.matpr.2021.07.281

Sarkar, D., Bali, R., & Sharma, T. (2017). *Practical Machine Learning with Python: A Problem-Solver's Guide to Building Real-World Intelligent Systems* (1st ed. edition). Apress.

Satia, G. A. W., Firmansyah, E., & Umami, A. (2022). Perancangan Sistem Identifikasi Penyakit pada Daun Kelapa Sawit (Elaeis guineensis Jacq.) dengan Algoritma Deep Learning Convolutional Neural Networks. *Jurnal Ilmiah Pertanian*, *19*(1), 1–10. https://doi.org/10.31849/jip.v19i1.9556

Singh, C., Walia, E., & Kaur, K. P. (2018). Enhancing color image retrieval performance with feature fusion and non-linear support vector machine classifier. *Optik*, *158*, 127–141. https://doi.org/10.1016/j.ijleo.2017.11.202

Wang, Z., Ruan, Q., & An, G. (2016). Facial expression recognition using sparse local Fisher discriminant analysis. *Neurocomputing*, *174*, 756–766. https://doi.org/10.1016/j.neucom.2015.09.083

Wen, T., Yan, J., Huang, D., Lu, K., Deng, C., Zeng, T., Yu, S., & He, Z. (2018). Feature Extraction of Electronic Nose Signals Using QPSO-Based Multiple KFDA Signal Processing. *Sensors (Basel, Switzerland)*, *18*(2), 388. https://doi.org/10.3390/s18020388

Zafeiriou, S. (2012). Subspace Learning in Krein Spaces: Complete Kernel Fisher Discriminant Analysis with Indefinite Kernels. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, & C. Schmid (Eds.), *Computer Vision – ECCV 2012* (pp. 488–501). Springer. https://doi.org/10.1007/978-3-642-33765-9_35

*name of corresponding author

e- ISSN : 2541-2019
p- ISSN : 2541-044X

*name of corresponding author