

# The effect of features combination on coloscopy images of cervical cancer using the support vector machine method

Retno Supriyanti<sup>1</sup>, Andreas Sahir Aryanto<sup>1</sup>, Mohammad Irham Akbar<sup>1</sup>, Eman Sutrisna<sup>2</sup>,  
Muhammad Alqaaf<sup>3</sup>

<sup>1</sup>Department of Electrical Engineering, Faculty of Engineering, Jenderal Soedirman University, Purwokerto, Indonesia

<sup>2</sup>Department of Medical, Faculty of Medicine, Jenderal Soedirman University, Purwokerto, Indonesia

<sup>3</sup>School of Information Science, Nara Institute of Science and Technology, Ikoma, Japan

## Article Info

### Article history:

Received Nov 24, 2023

Revised Feb 15, 2024

Accepted Feb 28, 2024

### Keywords:

Cervical cancer

Colposcopy

Early detection

Feature combination

Support vector machine

## ABSTRACT

Cervical cancer is cancer that grows in cells in the cervix. This cancer generally develops slowly and only shows symptoms when it has entered an advanced stage. Therefore, it is crucial to detect cervical cancer early before serious complications arise. One way to detect cervical cancer early is to use colposcopy, which is to look closely at the condition of the cervix to find changes in cells in the cervix that have the potential to become cancer. However, this method requires the expertise of an obstetrician. This research proposes the use of image processing techniques to create automatic early detection of cervical cancer based on coloscopy images. In this paper, we will discuss image selection using an approach in the form of comparing the weights of feature vectors and then using a data distribution threshold, features that are not too influential can be eliminated. Image classification uses the support vector machine (SVM) method, which makes it possible to distinguish normal images from abnormal images. Classification with feature selection and merging results can improve the consistency of SVM model performance evenly across all four SVM kernels.

*This is an open access article under the [CC BY-SA](#) license.*



## Corresponding Author:

Retno Supriyanti

Department of Electrical Engineering, Faculty of Engineering, Jenderal Soedirman University

St. HR. Boenjamin 708 Purwokerto, Central Java, Indonesia

Email: retno\_supriyanti@unsoed.ac.id

## 1. INTRODUCTION

Cervical cancer is a significant health problem throughout the world. Data in 2020 shows that around 600,000 women have cervical cancer worldwide and more than 340,000 cases of death from its consequences [1]. In addition, the fact that cervical cancer is ranked second as a cancer in the female reproductive system makes it a severe challenge in the world of global health. As a result of this reality, cervical cancer also occupies the fourth position as the deadliest disease for women throughout the world [2]. Apart from significant health losses, cervical cancer also has implications for widespread economic and social losses. One effective way to overcome the burden that can be caused by cervical cancer is to carry out early detection or pre-cancer detection. The pre-cancer detection process has a vital role in the prevention and treatment of this disease. Detection of pre-cancer is critical, considering the slow nature of cervical cancer in its development [3]. Effective pre-cancer detection aims to identify lesions on the cervix for medical treatment before they develop into more severe cancer. Obstacles such as limited resources and access to health facilities, as well as a lack of awareness regarding the importance of early detection, are the main obstacles in treating cervical cancer. One of the significant pre-cancer detection techniques is the colposcopy technique. Colposcopy is a diagnostic procedure that can identify lesions or abnormal tissue on the cervix

and uterine wall [4]. Colposcopy allows for observation at up to tenfold magnification, making it an essential instrument in identifying such lesions or abnormal tissue [5]. Based on this, it is possible to use colposcopy images in the early detection of cervical cancer.

Several studies have used coloscopy images to detect cervical cancer. Kalbhor and Shinde [6] created a system framework that will be used to detect cervical pre-cancer based on colposcopy images with the ultimate aim of having high effectiveness compared to existing methods. The researcher uses features in the form of gray-level co-occurrence matrix (GLCM), gray-level run length matrix (GLRLM), and histogram of gradients (HOG), which will then be combined to continue with classification. The researcher used six classification algorithms: naïve Bayes, Bayes Net, random tree, random forest, decision table, and logistic. Amalia *et al.* [7] are applied and compared to two edge detection operators, isotropic and canny, in processing coloscopy images to diagnose cervical cancer. Rahaman *et al.* [8] reviewed the development of optical techniques in the diagnosis of cervical cancer. Xue *et al.* [9] conducted a meta-analysis to assess the performance of using deep learning in four subgroups, namely cancer type (breast or cervical), validation type (internal or external), imaging modality (mammography, ultrasound, cytology, or colposcopy), and deep learning algorithm versus doctor.

Origoni *et al.* [10] evaluated the accuracy of the colposcopy method in diagnosing cervical cancer through the results of quality control and quality assurance assessments in a survey of academic and teaching hospitals in Italy. Tawalbeh *et al.* [11] compared six feature-combining techniques: canonical correlation analysis, discriminant correlation analysis, and most minor absolute reduction and selection-operators, independent component analysis, principal component analysis, and combining. Srinivasan *et al.* [12] proposed a cervical cancer detection method using cervigram images. In this method, they used a histogram to refine the contour of cervical cancer images and then made limited transformations to produce images with multiple resolutions. In their classification, they use forward propagation and back propagation neural networks. Shamsunder *et al.* [13] researched to validate artificial intelligence in conducting image assessments as a screening for uterine health. Park *et al.* [14] used a deep learning method based on cervical area segmentation to extract coloscopy image features automatically. Qin *et al.* [15] performed a systematic review and meta-analysis on the performance of colposcopy in diagnosing cervical intraepithelial neoplasia.

Li *et al.* [16] conducted a survey from February 2020 to December 2022 on women regarding colposcopy. Mustafa *et al.* [17] reviewed cervical cancer detection techniques chronologically, from the traditional Pap smear test to the latest computer-aided detection (CAD) systems. Yan *et al.* [18] conducted a prospective study in a gynecology clinic from March to September 2021. They recruited women with cervical abnormalities. Then, they evaluated the diagnostic performance of optical coherence tomography (OCT) in detecting these abnormalities. Chen *et al.* [19] used artificial intelligence in the recognition and classification of colposcopic images of cervical cancer. Nie *et al.* [20] proposed a new specular reflection detection and removal method for endoscopic images based on brightness level classification. It can effectively detect specular areas in endoscopic images with different brightness levels, thereby improving the operating efficiency of the algorithm. Al-Jabbar *et al.* [21] developed three techniques for diagnosing breast cancer. The first technique is breast cancer diagnosis using an artificial neural network modified with several features. From this research to the best of our knowledge, no research has discussed the effect of combining features in coloscopy images based on the support vector machine (SVM) method in the early detection of cervical cancer. In our previous research [22]–[26], We use many image processing techniques to carry out artificial intelligence-based classification for early detection of several diseases using images from various modalities. In this paper, we discuss the use of coloscopy images based on feature fusion in improving the performance of the SVM method in cervical cancer classification.

## 2. PROPOSED METHOD

The feature extraction and selection stage is a novelty in this research. At this stage, we calculate color and texture moment features from colposcopy images that have undergone initial processing. The feature extraction results will be divided into two feature matrices, which will be used as input vectors in SVM classification. The weight values will be calculated for both models as an approach to feature selection. The weight value will be compared with the threshold value in the form of the skewness threshold or direction of data distribution, namely the mean and median. When the weight value of a feature is smaller than the threshold value, the feature will be eliminated. So, features with a weight value more significant than the threshold value are produced. The feature selection results will be combined into a matrix, which will be classified using SVM. At this classification stage, it is possible to design a machine learning classification model using the SVM algorithm based on colposcopy image features that have been previously selected. The classification model will be used as the final and then compared with the two models before selecting and combining features. The use of SVM as a pattern recognition algorithm or machine learning is based on classification objectives; namely, the classification in this research is binary. The SVM algorithm is very

suitable and easy to carry out binary classification. Also, it can use a kernel to accept input vectors of up to  $n$  features.

### 3. METHOD

#### 3.1. Input images

This research used acetic acid visual inspection or *inspeksi visual asam asetat* (IVA) colposcopy images sourced from the International Agency for Research on Cancer (IARC) colposcopy image bank, with 161 images (70 abnormal images and 91 standard images) [27]. Colposcopy is a pre-cancer detection procedure that aims to find out whether there are indications of abnormal cells in the cervix. These abnormal cells generally occur at the opening of the cervix to the uterine canal. This procedure allows medical experts to open and view or photograph the cervix up to the uterine wall [28]. IVA is a 'naked eye' examination method to observe pre-cancer detection procedures, namely colposcopy, by identifying the size of the lesion before it becomes invasive and invasive. This method allows using an acetic acid solution of around 3%-5% on cotton wool and then applying it to the cervical tissue during colposcopy. The photographing or observation process requires assistance in the form of a halogen lamp. This process also takes about one minute to get the desired results. The results of an IVA are positive (indicating a lesion) if there is aceto whitening in the tissue or cervical cell tissue with white-like spots. It will have a negative value if there is no aceto whitening. This occurs due to increased nuclear protein and cytokeratin in the cervical epithelial tissue, causing striking color differences. Then, the squamous epithelial tissue on the cervix that does not have lesions will be pink [29]. Figure 1 shows examples of some colposcopy images used in this research.



Figure 1. Examples of several colposcopy images

#### 3.2. Feature extraction and selection

Feature extraction is converting an image into a group of features. This feature group is a number that describes certain information from an image. There are various groups of features, one of which is color and texture moments. Color moment can be interpreted as a scale that can differentiate images based on color features. The color moments in the image are interpreted as pixel probability distributions. There are three components in color moments: mean, skewness, and standard deviation. Texture is the most crucial feature for various images that appear everywhere in nature, such as medical images and sensor images. Texture is the surface manifestation of the human visual system towards natural objects. In this research, the GLCM texture will be used, which consists of entropy, contrast, correlation, energy, and homogeneity [30].

#### 3.3. Support vector machine

Machine learning is an interdisciplinary field built on ideas from cognitive science, computer science, statistics, and other mathematical sciences. In machine learning, classification is a method for approaching (supervised learning) a given data set by creating a model that can separate the data into the desired number [31]. Many classification techniques already exist and have been developed to date, one of which is SVM. Based on pattern recognition in classification, creating a model that can provide appropriate input and output pairs is possible. SVM is a classification and regression technique that maximizes prediction accuracy. SVM works by mapping data into a high-dimensional feature space so that the data can be categorized even if the data cannot be separated linearly. The separation between these data classes is called a hyperplane [32].

#### 4. RESULTS AND DISCUSSION

Colposcopy is a pre-cancer detection procedure that aims to find out whether there are indications of abnormal cells in the cervix. These abnormal cells generally occur at the opening of the cervix to the uterus. This procedure allows medical experts to open and view or photograph the cervix up to the uterine wall. The research begins by cropping the colposcopic image, aiming to reduce noise in the image. This process will be done manually with region of interest (ROI). After cutting, the image will be divided into three data groups for further processing: test, training, and validation. The data sharing ratio is 80:20 between training data and test data, referring to the paper written by Gholamy *et al.* [33]. Figure 2 shows the results of image cropping that will be used in the next stage of analysis. Next, the image will be converted into gray, as shown in Figure 3. Meanwhile, Table 1 shows the division of image data as training, test, and validation data.



Figure 2. Result of cropping image

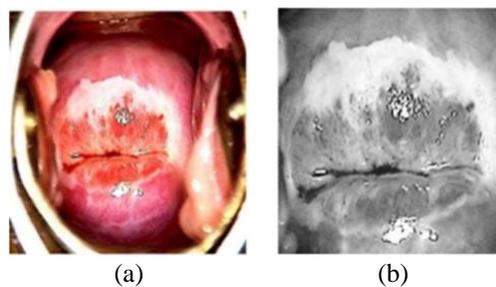


Figure 3. Result of pre-processing image: (a) original image and (b) the result of pre-processing

Table 1. Data distribution

Image	Total	Train	Validation	Test
Normal	91	59	14	18
Abnormal	70	45	11	14

This feature extraction will be carried out simultaneously for color moment features (mean, skewness, and standard deviation) and texture features (contrast, correlation, energy, homogeneity, and entropy). In this experiment we used the Python programming language tool. We use the numpy library to calculate the mean of the pixels in the image. Then, using the np. Float data type increases the accuracy of the calculation. The image must be created in a 1-dimensional array to be calculated. The skewness calculation must be ensured only in the image array column, so there is a parameter axis = 0. We use the numpy library to calculate the variation of pixels in the image. Because the image is a population, not a sample, because all image pixels are segmented, the parameter ddof = 0 is needed. Then, do the square root of the variation value. We use the scikit-learn library to calculate the entropy of pixels in the image. In contrast calculations, several parameters need to be set; i) distance will select how many pixels will be the reference in the calculation (set 1 and 2 pixels), ii) angles will set the angle of the second pixel with the first reference pixel for calculations (select 0, 45, 90, and 135 degrees), and iii) levels will set the type of image used because the image is grayscale; it will have a value of 256. This parameter will be used for all texture feature calculations. The calculation result will be in the form of an array of more than one dimension. Therefore, it will be created in 1 dimension first, and only the value of the first element will be taken.

The classification model uses a SVM as the algorithm. The SVM model is trained using features from the training data, then its hyperparameters are tuned or calibrated using validation data and then tested with "unseen data," namely test data. At this stage, classification is carried out for both groups of features.

The system we developed can access files containing training data for the model. The data consists of features and corresponding labels or classes. These features will be retrieved and stored in two variables, namely 'x' for features and 'y' for labels. Next, the program will prepare a list of parameters used in the model-tuning process. These parameters include the 'C' parameter, 'gamma,' and the kernel types that can be used, such as 'linear,' 'polynomial,' 'radial basis function (RBF),' and 'sigmoid.' Next, using the GridSearchCV function will allow tuning of the combinations in the list of parameters that have been set. GridSearchCV will perform parameter tuning by cross-validating ten times for each existing combination. This is done to avoid overfitting and test the stability of the model. The best parameter results will automatically be included as hyperparameters in the model. After obtaining appropriate hyperparameters, the model performance will be calculated using accuracy, sensitivity (recall), precision, F1 score, and specificity. The results of model performance calculations are also visualized in a confusion matrix to make analysis easier. Figure 4 shows the performance graph of the SVM model with several color feature-based kernels.

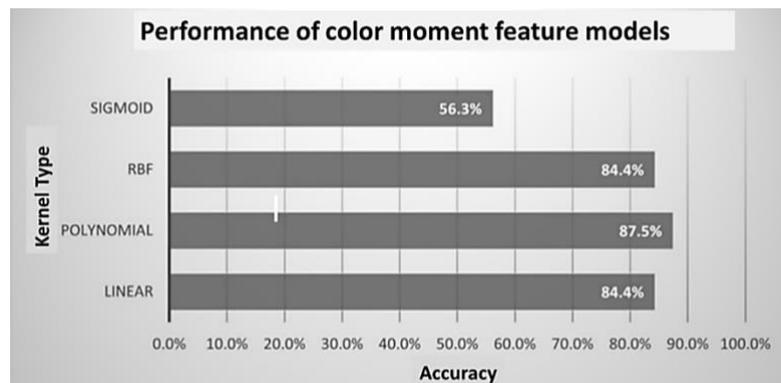


Figure 4. Accuracy graph of the SVM model of color moment features

As in the previous stage, this stage uses texture feature data: contrast, correlation, energy, homogeneity, and entropy. The texture feature matrix is used as input data. Four kernels were also used in making this classification model, and three experiments were carried out for each kernel. Then, the model's performance results are directly calculated and shown in Figure 5.

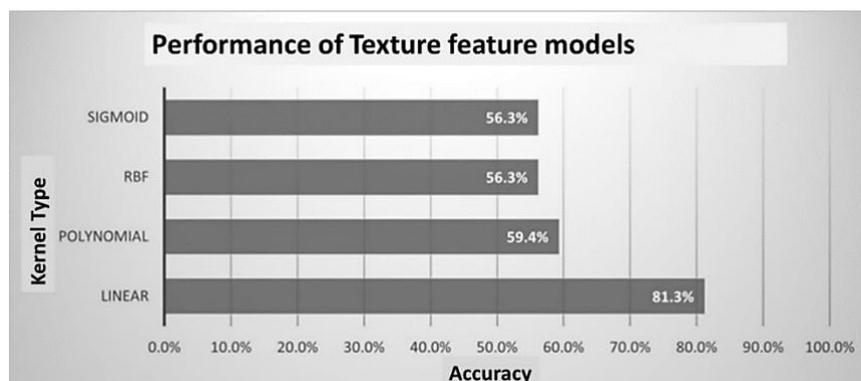


Figure 5. Graph of texture feature SVM model accuracy

After carrying out the model creation process and calculating the performance of each model, it is then analyzed how the features from each feature group play a role in helping the model classify 'abnormal' and 'normal' images. This analysis can be carried out using the weight values of the linear kernel for each feature vector. This approach can also be called feature importance or the level of feature importance. An analytical approach for feature selection using weight values in a linear kernel for each feature vector can be used to combine and select features. The model is trained using a linear kernel to access the linear SVM

weights. These weights can be accessed using the coef [0] property, namely accessing the correct weights in distinguishing class 0 (normal) and not 0 (abnormal). The value used is only the absolute (positive) value of the weight, which is then sorted from largest to smallest value and then sorted according to the features in the input feature matrix. The results of this features importance approach can be seen in Figures 6 and 7.

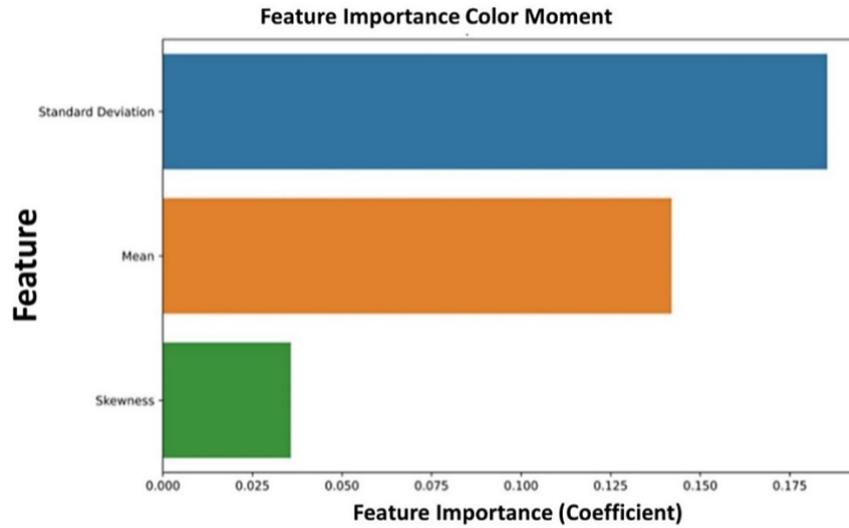


Figure 6. Visualization of features importance of color moments

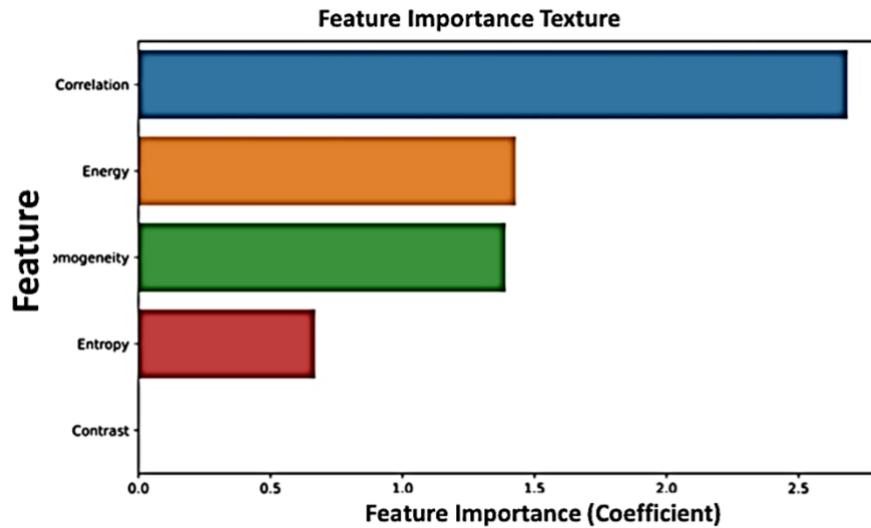


Figure 7. Visualization of features importance of texture

Next, using the weight value of each feature group, the threshold value can be found to decide which features will be eliminated. The threshold parameters used are based on the distribution of the weight data in Figures 6 and 7. The parameters that describe the data distribution are the mean and median. To choose a suitable parameter as a threshold value, we analyze to see the skewness of the data. Data skewness can be seen by comparing the mean and median values of the weights of the two feature groups. The equations for the mean and median are shown in (1) and (2):

$$\bar{x} = \frac{x_1+x_2+x_3+...+x_n}{n} \tag{1}$$

$$median = \frac{n+1}{2} \tag{2}$$

With;  $\bar{x}$  = mean,  $x_n$  = data to - n,  $n$ = number of data.

The mean and median values for the feature weight data in the two feature groups are obtained, as shown in Table 2. Referring to Table 2 in the color moment feature, the mean and median values are not too far apart, indicating that the data is distributed symmetrically so that both thresholds can be used. Meanwhile, in the texture feature weight data, the average value is far below the median value, which indicates that the data is skewed towards positive values, so using the median as the threshold is much better. Therefore, a threshold value can be used from the median value of each feature group. The selected feature must have a value greater than or equal to the median value so that the feature selection results can be seen in Table 3.

Table 2. The mean and median feature weights

Feature	Mean	Median
Color moment	0.1204	0.1410
Texture	1.035	1.390

Table 3. Feature selection results

Feature	Selected	Eliminated
Color moment	Mean dan standard deviation	skewness
Texture	Energy, homogeneity, dan correlation	entropy dan contrast

After obtaining the selected features based on their weight values, these features can be combined into one matrix, which can be used to create models and calculate their performance. When classifying using SVM, we use combined feature data: mean, standard deviation, energy, homogeneity, and correlation. In creating the classification model, four kernels were used, and three experiments were carried out for each kernel. Figure 8 shows the accuracy results for the four kernels.

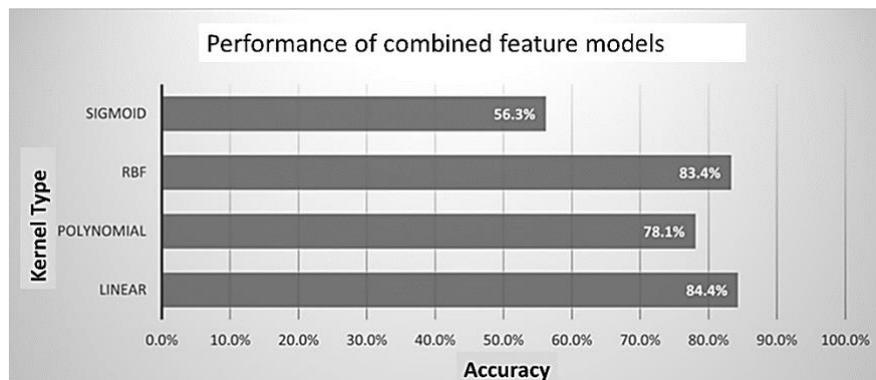


Figure 8. Accuracy graph of combined feature SVM model

### 5. CONCLUSION

In this research, we apply feature selection with an approach that compares the weights of feature vectors on the linear hyperplane SVM. This modification makes it possible to get the most relevant features from the color and texture moment features. The research results show that combining selected features influences the SVM model's performance. The SVM model's performance is improved considerably by using only texture features. However, performance in the sigmoid kernel remains the same. In the final results, the linear kernel has the best performance. This case happens because the linear kernel has the most straightforward equation: just the dot product of the weights and feature vectors. So, it can be interpreted that image feature data can be separated linearly. A pre-cancer detection system with the best performance was successfully created using a combination of selected features. This system improves performance consistency across four SVM kernels.

### ACKNOWLEDGEMENTS

This research was funded by BLU Jenderal Soedirman University through Fundamental Research Scheme. The author also thanks the International Agency for Research on Cancer (IARC) for permission to use the colposcopy image database.

## REFERENCES

- [1] H. Sung *et al.*, “Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021, doi: 10.3322/caac.21660.
- [2] L. Bruni *et al.*, “Human papillomavirus and related diseases in the world,” *ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre)*, Barcelona, Spain, 2023.
- [3] M. F. Janicek and H. E. Averette, “Cervical cancer: prevention, diagnosis, and therapeutics,” *CA: A Cancer Journal for Clinicians*, vol. 51, no. 2, pp. 92–114, Mar. 2001, doi: 10.3322/canjclin.51.2.92.
- [4] D. B. Cooper and C. J. Dunton, *Colposcopy*, Treasure Island, USA: StatPearls Publishing, 2024.
- [5] M. Kyrgiou *et al.*, “The up-to-date evidence on colposcopy practice and treatment of cervical intraepithelial neoplasia: The cochrane colposcopy & cervical cytopathology collaborative group (C5 group) approach,” *Cancer Treatment Reviews*, vol. 32, no. 7, pp. 516–523, Nov. 2006, doi: 10.1016/j.ctrv.2006.07.008.
- [6] M. Kalbhor and S. Shinde, “ColpoClassifier: a hybrid framework for classification of the cervigrams,” *Diagnostics*, vol. 13, no. 6, Mar. 2023, doi: 10.3390/diagnostics13061103.
- [7] A. Amalia, D. Rachmawati, I. A. Lestari, and C. Mourisa, “Implementation of canny and isotropic operator with power law transformation to identify cervical cancer,” in *Journal of Physics: Conference Series*, 2018, pp. 1–8, doi: 10.1088/1742-6596/979/1/012031.
- [8] A. Rahaman, A. Anantharaju, K. Jeyachandran, R. Manideep, and U. M. Pal, “Optical imaging for early detection of cervical cancer: state of the art and perspectives,” *Journal of Biomedical Optics*, vol. 28, no. 08, 2023, doi: 10.1117/1.jbo.28.8.080902.
- [9] P. Xue *et al.*, “Deep learning in image-based breast and cervical cancer detection: a systematic review and meta-analysis,” *npj Digital Medicine*, vol. 5, no. 1, 2022, doi: 10.1038/s41746-022-00559-z.
- [10] M. Origoni *et al.*, “Colposcopy accuracy and diagnostic performance: a quality control and quality assurance survey in Italian tertiary-level teaching and academic institutions—the Italian society of colposcopy and cervico-vaginal pathology (SICPCV),” *Diagnostics*, vol. 13, no. 11, 2023, doi: 10.3390/diagnostics13111906.
- [11] S. Tawalbeh, H. Alquran, and M. Alsalatie, “Deep feature engineering in colposcopy image recognition: a comparative study,” *Bioengineering*, vol. 10, no. 1, 2023, doi: 10.3390/bioengineering10010105.
- [12] S. Srinivasan, A. B. K. Raju, S. K. Mathivanan, P. Jayagopal, J. C. Babu, and A. K. Sahu, “Local-ternary-pattern-based associated histogram equalization technique for cervical cancer detection,” *Diagnostics*, vol. 13, no. 3, pp. 1–16, 2023, doi: 10.3390/diagnostics13030548.
- [13] S. Shamsunder, A. Mishra, A. Kumar, and S. Kolte, “Automated Assessment of digital images of uterine cervix captured using transvaginal device—a pilot study,” *Diagnostics*, vol. 13, no. 19, pp. 1–11, 2023, doi: 10.3390/diagnostics13193085.
- [14] J. Park, H. Yang, H. J. Roh, W. Jung, and G. J. Jang, “Encoder-weighted W-Net for Unsupervised segmentation of cervix region in colposcopy images,” *Cancers*, vol. 14, no. 14, 2022, doi: 10.3390/cancers14143400.
- [15] D. Qin *et al.*, “Colposcopic accuracy in diagnosing squamous intraepithelial lesions: a systematic review and meta-analysis of the international federation of cervical pathology and colposcopy 2011 terminology,” *BMC Cancer*, vol. 23, no. 1, pp. 1–17, 2023, doi: 10.1186/s12885-023-10648-1.
- [16] X. Li *et al.*, “Detection of cervical high-grade squamous intraepithelial lesions and assessing diagnostic performance of colposcopy among women with oncogenic HPV,” *BMC Women’s Health*, vol. 23, no. 1, pp. 1–9, 2023, doi: 10.1186/s12905-023-02538-2.
- [17] W. A. Mustafa, S. Ismail, F. S. Mokhtar, H. Alquran, and Y. Al-Issa, “Cervical cancer detection techniques: a chronological review,” *Diagnostics*, vol. 13, no. 10, May 2023, doi: 10.3390/diagnostics13101763.
- [18] L. Yan *et al.*, “Efficacy of optical coherence tomography in the triage of women with minor abnormal cervical cytology before colposcopy,” *PLoS ONE*, vol. 18, no. 3, pp. 1–19, 2023, doi: 10.1371/journal.pone.0282833.
- [19] X. Chen *et al.*, “Application of efficientNet-B0 and GRU-based deep learning on classifying the colposcopy diagnosis of precancerous cervical lesions,” *Cancer Medicine*, vol. 12, no. 7, pp. 8690–8699, Apr. 2023, doi: 10.1002/cam4.5581.
- [20] C. Nie, C. Xu, Z. Li, L. Chu, and Y. Hu, “Specular reflections detection and removal for endoscopic images based on brightness classification,” *Sensors*, vol. 23, no. 2, 2023, doi: 10.3390/s23020974.
- [21] M. Al-Jabbar, M. Alshahrani, E. M. Senan, and I. A. Ahmed, “Analyzing histological images using hybrid techniques for early detection of multi-class breast cancer based on fusion features of CNN and handcrafted,” *Diagnostics*, vol. 13, no. 10, 2023, doi: 10.3390/diagnostics13101753.
- [22] R. Supriyanti, M. R. Kurniawan, Y. Ramadhani, and H. B. Widodo, “Calculating the area of white spots on the lungs of patients with COVID-19 using the Sauvola thresholding method,” *International Journal of Electrical and Computer Engineering*, vol. 13, no. 1, pp. 315–324, 2023, doi: 10.11591/ijece.v13i1.pp315-324.
- [23] R. Supriyanti, F. F. R. Wibowo, Y. Ramadhani, and H. B. Widodo, “Comparison of conventional edge detection methods performance in lung segmentation of COVID19 patients,” *AIP Conference Proceedings*, vol. 2482, 2023, doi: 10.1063/5.0111683.
- [24] R. Supriyanti, M. Alqaaf, Y. Ramadhani, and H. B. Widodo, “Morphological characteristics of X-ray thorax images of COVID-19 patients using the Bradley thresholding segmentation,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 24, no. 2, pp. 1074–1083, 2021, doi: 10.11591/ijeecs.v24.i2.pp1074-1083.
- [25] R. F. W. Supriyanti, Y. Ramadhani, and H. B. Widodo, “Comparison of conventional edge detection methods performance in lung segmentation of COVID19 patients,” *AIP Conference Proceedings*, vol. 2482, no. 1, 2023, doi: 10.1063/5.0111683.
- [26] R. Supriyanti, A. R. Subhi, E. J. Ashari, F. Ahmad, Y. Ramadhani, and H. B. Widodo, “Simple classification of the alzheimer’s severity in supporting strengthening the diagnosis of patients based on ROC diagram,” in *IOP Conference Series: Materials Science and Engineering*, pp. 1–7, Dec. 2020, doi: 10.1088/1757-899X/982/1/012007.
- [27] International Agency for Research on Cancer, “Cancer screening at IARC,” *World Health Organization*, 2024. [Online]. Available: <https://screening.iarc.fr/>
- [28] “Colposcopy,” *Mayo Clinic*, 2024. [Online]. Available: <https://www.mayoclinic.org/tests-procedures/colposcopy/about/pac-20385036>
- [29] K. C. Carr and J. W. Sellors, “Cervical cancer screening in low resource settings using visual inspection with acetic acid,” *Journal of Midwifery and Women’s Health*, vol. 49, no. 4, pp. 329–337, 2004, doi: 10.1016/j.jmwh.2004.02.015.
- [30] E. A. B. D. Silva and G. V. Mendonça, “Digital image processing,” in *The Electrical Engineering Handbook*, New Jersey: Elsevier, 2005, pp. 891–910, doi: 10.1016/B978-012170960-0/50064-5.
- [31] J. M. Luna, “Introduction to data mining,” in *Periodic Pattern Mining*, Singapore: Springer, 2021, pp. 1–22, doi: 10.1007/978-981-16-3964-7\_1.
- [32] J. Cervantes, F. G. -Lamont, L. R. -Mazahua, and A. Lopez, “A comprehensive survey on support vector machine classification: applications, challenges and trends,” *Neurocomputing*, vol. 408, pp. 189–215, 2020, doi: 10.1016/j.neucom.2019.10.118.
- [33] A. Gholamy, V. Kreinovich, and O. Kosheleva, “Why 70/30 Or 80/20 relation between training and testing sets: a pedagogical explanation,” *Departmental Technical Reports (CS)*, vol. 1209, pp. 1–6, 2018.

**BIOGRAPHIES OF AUTHORS**

**Retno Supriyanti**    is a professor at Department of Electrical Engineering, Jenderal Soedirman University, Indonesia. She received her Ph.D. in March 2010 from Nara Institute of Science and Technology, Japan. Also, she received her M.S. degree and bachelor degree in 2001 and 1998, respectively, from Department of Electrical Engineering, Gadjah Mada University, Indonesia. Her research interests include image processing, computer vision, pattern recognition, biomedical application, e-health, tele-health, and telemedicine. She can be contacted at email: [retno\\_supriyanti@unsoed.ac.id](mailto:retno_supriyanti@unsoed.ac.id).



**Andreas Sahir Aryanto**    is a student at Department of Electrical Engineering, Jenderal Soedirman University, Indonesia. His research interest image processing field. He can be contacted at email: [andreas.aryanto@mhs.unsoed.ac.id](mailto:andreas.aryanto@mhs.unsoed.ac.id).



**Mohammad Irham Akbar**    is an academic staff at Department of Electrical Engineering, Jenderal Soedirman University, Indonesia. He received his M.S. from Gadjah Mada University, Indonesia, and his bachelor degree from Jenderal Soedirman University, Indonesia. His research interest including computer network, decision support system, data mining, and medical imaging. He can be contacted at email: [mohammad.irham@unsoed.ac.id](mailto:mohammad.irham@unsoed.ac.id).



**Eman Sutrisna**    is an academic staff at Faculty of Medical, Jenderal Soedirman University, Indonesia. He received his Ph.D. and M.S. degree from Padjajaran University, Indonesia. Also, bachelor degree from Sebelas Maret University, Indonesia. His research interest including public health, biomedical, and pharmacology. He can be contacted at email: [eman.sutisna@unsoed.ac.id](mailto:eman.sutisna@unsoed.ac.id).



**Muhammad Alqaaf**    received his bachelor degree from Department of Electrical Engineering, Jenderal Soedirman University, Indonesia; M.S. degree in Nara Institute of Science and Technology, Japan. Currently he is a doctor student in Nara Institute of Science and Technology, Japan. His research interest image processing field and bioinformatics. He can be contacted at email: [muhammad.alqaaf\\_subandoko.mb5@is.naist.jp](mailto:muhammad.alqaaf_subandoko.mb5@is.naist.jp).