

CNN and SVM Based Classifier Comparison to Detect Lung Nodule In Computed Tomography Images

I Wayan Budi Sentana¹, Sri Andriati Asri²

Department of Information System
Bali State Polytechnic
Bali, Indonesia

¹iw.budi.s@gmail.com, ²andriati_s@yahoo.com

Naser Jawas³

Computer Engineering Department
STM IK STIKOM Bali
Denpasar, Indonesia

³naser.jawas@gmail.com

Anggun Esti Wardani⁴

Radiology Unit, Airlangga University Education Hospital
Airlangga University
Surabaya, Indonesia

⁴angioanggun@gmail.com

Abstract—Convolutional Neural Networks (CNN) are natural based classification algorithm that combine Multiple Layer Perceptron (MLPs). Meanwhile, support vector machines (SVM) is a mathematical-based classification algorithm that naturally have supervised learning models. In some research related to image processing, each algorithm has its owned supremacy as well as the drawback. None of the previous studies compare both algorithm when they are utilized to detect nodule located in the pulmonary or lung images produced by Computed Tomography (CT) scan. Hence, this research comparing the two algorithms in case of lung nodule detection in CT images, since detecting lung nodule in CT images is still challenging. SVM-based classifier is preceded by feature extraction as its common behavior of mathematical based classifier. There are three algorithms use to conduct feature extraction process, namely Hu moment invariant, Haralick and Color Histogram extraction. In the opposite, CNN-based classifier consists of three layers convolution for training and testing steps. The result shows that SVM has better results than CNN in case of computing speed. Meanwhile have a better accuracy in detecting lung nodule. The results of the test analysis show that the extractor feature when preprocessing conduct before being classified by SVM makes the computing process faster. The accuracy of SVM-based classifier can be improved by adjusting some computation variables in feature extraction stages, such as adding more bins in the color histogram extraction. Those adjustment will lead to more computation times.

Keywords—CNN; SVM; lung nodule; computed tomography images

I. INTRODUCTION

Pulmonary Nodule is one of the parameter that use to diagnose the appearance of certain pulmonary disease such as emphysema, bronchitis, tuberculosis and also lung cancer [1]. In medical image such as Computed Tomography or Magnetic Resonance Imager, it's appeared in the thoracic area as white

patches of lumps. According to Report by [2], [3], lung cancer becomes one of the deadliest diseases that cause 1.2-1.56 million deaths annually. It is predicted that those number will continue to increase and will reach 17 million lung cancer case by 2030 [4]. From the whole lung cancer cases, 80% of the cases are identified after medium stage [4], which means that it will more difficult to handle by medical treatment. Hence, early detection by identifying the appearance of lung nodule will be able to provide better handling to the patient.

Various techniques of radiology has been utilized to capture thoracic internal organ in order to identify lung nodule, such as Computed Tomography, Rontgen imager, and Magnetic Resonance Imager [5], [6]. Most of the Rontgen images are suitable for prior screening stage only. Meanwhile, to have a detail view of nodule position, most of the radiologist will utilize CT or MRI[1]. In certain case, radiologist utilized MRI to have a better image results, but MRI has an expensive operational cost and has a drawback in soft tissue organ as well as lungs. Therefore, the CT scan image are still become an option for use in medical practices and researches. Although it involved in so many researches, detecting lung nodule in CT scan images is till challenging to conduct.

In the other hand, there so many researches utilized Convolutional Neural Network (CNN)-based algorithm have shown their own superiority in CT scans image processing. Some of them are shown by [7-10]. In the opposite, some researches in SVMs also shows the same, as explored by [11], [12] and [13]. However, none of those researches compared the performance of both CNN and SVM algorithms to classify the lung nodule cases. Hence, this research compares those algorithms in case of accuracy and computation times.

II. RELATED RESEARCH

A. Hu's Moment Invariants

This concept of moment invariants is derivative of geometrical analytic proposed by Sylvester and Cayley [14]. Hu proposed the concept of algebra moment invariants that combine several general algebra moments to form a group of algebra moments. This group of moments, or better known as Hu moments, are invariance in object's scale, translation and rotation change [14]. Regarding the two dimensional geometric moment in digital image, a density distribution function $f(x, y)$ of order $p + q$ are elaborated as

$$m_{pq} = \int_{x,y \in C} \int x^p y^q f(x, y) dx dy \quad (1)$$

where $p, q = 0, 1, 2, \dots, n$

and those value are not considered as invariant. The integrals are represented over the total pixel-area of the digital image object including its boundary that implies computational complexity of order $O(N^2)$. To deliver the formula in discrete function, $f(x, y)$, represented in a discrete points domain of $M \times N$, so the equation (1) the represented in the form.

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q f(x, y) \quad (2)$$

The order $p+q$ central moment is elaborated as followed

$$\mu_{pq} = \int_{x,y \in C} \int (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (3)$$

$$\bar{x} = \frac{m_{10}}{m_{00}} = \frac{\int_{x,y \in C} \int x f(x, y) dx dy}{\int_{x,y \in C} \int f(x, y) dx dy} \quad (4)$$

$$\bar{y} = \frac{m_{01}}{m_{00}} = \frac{\int_{x,y \in C} \int y f(x, y) dx dy}{\int_{x,y \in C} \int f(x, y) dx dy} \quad (5)$$

The scope of digital area or the object is presented by m_{00} . The feature of scale invariant also represented in central moment scaled of η_{pq} . The central moment after normalization of order $(p+q)$ is presented in formula(6).

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\frac{(1+p+q)}{2}}} \quad (6)$$

With $p + q \geq 2$. In addition, the translational invariance must be followed by only tailoring central moments directly.

B. Haralick Feature Extraction

Feature extraction is tailored to find unique features of the existing image. A concept that can be done is to find a feature based on color differences. A well-known method of extracting color-based features is Haralick, Haralick extracted the feature from color concurrence matrices. The main concept of this algorithm is to distributed the value of neighboring color pixel of a matrix in a certain size. Color concurrence matrices are statistical features that measuring distribution of color in an image and recompute the value based on the spatial interactions among the value of pixels[15]. In each color space, these matrices are defined (C_1, C_2, C_3) of figure 1. Let C_k and $C_{k'}$, be two of the three color components of this space $k, k' \in \{1, 2, 3\}$ and $M^{C_k, C_{k'}}[I]$, the color concurrence matrix which measures the spatial interactions among the color components C_k and $C_{k'}$ of the pixels in the image I. The cell $M^{C_k, C_{k'}}[I](i, j)$

of this matrix consist of the number of times that a pixel P whose color component value $C_k(P)$ is equal to i , has in its neighborhood denoted N , a pixel Q whose color component $C_{k'}(Q)$ is equal to j . Figure 1 shows $N_N = 7$ different 3×3 neighborhoods N considered to compute the color concurrence matrices.

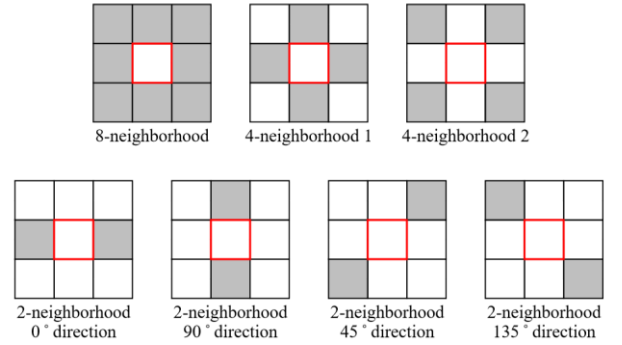


Fig. 1. 2×3 neighborhoods that labeled as grey [15]

For a given neighborhood, the image I can be characterized by $N_M = 6$ color co-occurrence matrices : $M^{C_1, C_2}[I]$, $M^{C_2, C_2}[I]$, $M^{C_3, C_3}[I]$, $M^{C_1, C_2}[I]$, $M^{C_1, C_3}[I]$ and $M^{C_2, C_3}[I]$. Since the matrices M^{C_2, C_1} , $M^{C_3, C_1}[I]$ and $M^{C_3, C_2}[I]$ are symmetric to the matrices $M^{C_1, C_2}[I]$, $M^{C_1, C_3}[I]$ and $M^{C_2, C_3}[I]$, they are not in used.

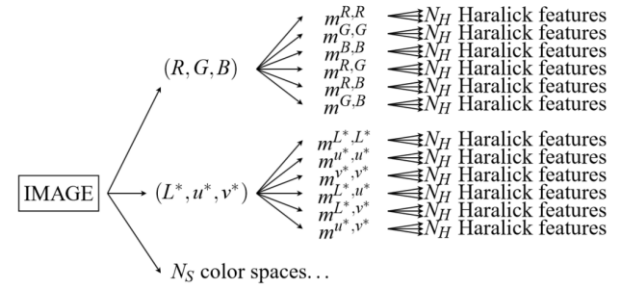


Fig. 2. Color texture features[15]

Since the algorithm compute local interaction among pixel's values, the color concurrence matrices become susceptible to striking differences of spatial resolution value. Hence, normalization is needed by these matrices to decrease this sensitivity. This normalization conducted by the total co-occurrence number $\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} M^{C_k, C_{k'}}[I](i, j)$ where N is the level of quantification number in the color components. The normalization result of color concurrence matrix $M^{C_k, C_{k'}}[I](i, j)$ is denoted by:

$$m^{C_k, C_{k'}}[I](i, j) = \frac{M^{C_k, C_{k'}}[I](i, j)}{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} M^{C_k, C_{k'}}[I](i, j)} \quad (7)$$

The color texture in the image are highly characterized by color concurrence matrices. However, the real image has a large amount of information that hard to be compute in color texture classification. Hence, to reduce the computation as well as preserving the relevance and important information of these descriptors, Haralick introduced the use of $N_H = 14$ features,

that denoted f_H^1 to f_H^{14} , which is resulted from each matrix [14]. As a consequence, an image then characterized by $N_f = N_M \times N_H \times N_S = 6 \times 14 \times 28 = 2352$ color texture features x^f , $f = 1, \dots, N_f$ as shown in figure 2.

C. Color Histogram

The next step after the feature extracted by Haralick is to find the threshold of the color feature. The color feature extraction is most widely used visual features in image retrieval as well as indexing[16]. The advantages of color feature are the ability to represent visual content of an image, have a simple color information extraction and have a high efficiency, relatively powerful in image separation, fairly robust to background complication and independent in accordance to image orientation and size. In this case a Histogram sum out the number of pixels based on each color group and create the total histogram by examined each pixel in an image and adding value to the appropriate bin of the histogram. In accordance to scale changes, translation, partial occlusion and axis rotation, Color histogram are subjected to be relatively invariant.

The dissimilarity are measure by the following distance formula:

$$d = \alpha(\alpha_1(\sum_{b=1}^L Wh_b \sum_{k=1}^K |h_b^I(k) - h_b^Q(k)|) + \beta_1(\sum_{k=1}^K |h^I(k) - h^Q(k)|)) + \beta(\sum_{b=1}^L Ws_b |S_b^I - S_b^Q|) \quad (8)$$

I represent the image that subjected to be extracted and Q represent the image query. h_b showing the histogram of b^{th} level of quadtree blocks for $b= 1,2,3 \dots,L$, and L represent the last level. Here L is 8 and h_b can be $h_1 = h_{128 \times 128}$, $h_2 = h_{64 \times 64}$, \dots , $h_8 = h_{1 \times 1}$. h_b and then normalized using the following:

$$h_b = \frac{h'_b}{M'_b \times N'_b \times S_b} \quad (9)$$

In which h'_b is the unnormalized histogram of b^{th} block level in the size of $M'_b \times N'_b$ and S_b represent the number of homogenous blocks in certain level. Wh_b represent the weight of h_b and selected the unnormalized as $Wh_b = (0,0.5,0.05,0.2,0.9,0.4,0)$ and this weight is normalized according to the existence of blocks as follow:

$$\text{sum_}Wh_b = \begin{cases} 0 & \text{if the have any blocks in } b^{\text{th}} \text{ level} \\ Wh_b & \text{if has one block/blocks in } b^{\text{th}} \text{ level} \end{cases}$$

$$Wh_b = \frac{Wh'_b}{\sum_{b=2}^{L=1} \text{sum_}Wh_b} \quad (10)$$

K represent the histogram bins number. Meanwhile, h represent the whole color histogram. Ws_b represent weight of S_b . α_1, β_1 denote relation combination of h_b with h and α, β denote relation of histogram complexity relation and $\alpha_1 + \beta_1 = 1, \alpha + \beta = 1$. Defined $\alpha_1 = \beta_1 = \alpha = \beta$.

D. SVM Classification

Classification algorithm needed to identify an object that subjected into certain classes. In this case SVM is tailored to

identify whether a feature extracted from an image is a nodule or not. As commonly classification algorithm, SVM have two stage which is training and testing. SVM basic concept is to define decision planes that isolated an object into certain group. A decision plane is used to separates between a group of objects that have different class memberships. In the case where the data can separate using linear line, the SVM hyperplane can effectively separate the objects into two classes. Meanwhile for the case that cannot separate using linear line or better known as linear inseparable, the value in original space need to be transform into feature space using a nonlinear transformation. To develop optimal hyperplane, SVM work with an iterative training algorithm, that is used to create minimum error function. A general formula for SVM classification are represented by mathematical optimization case as bellow[17][11][12]:

$$\min \Phi(w) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N \xi_i$$

Subject to

$$y_i (w^T \cdot \phi(x_i) + b) \geq 1 - \xi_i, (i = 1,2,3,\dots,N) \quad (11)$$

where C work as penalty function, and ξ_i is a parameter that handling inseparable data. Minimizing object function of $\Phi(w)$ in formula (12) are equal to Maximizing the distance between two classes as well as minimizing the misclassification rate. Parameter C use to control the trade-off between the class distance and slack variable penalty. The i index, labeling the N training object, $y_i \in \pm 1$ represents the labels of the class and x_i denoted independent variables. Kernel function denoted by ϕ is used in transformation of the data into feature space. There are three common kernel functions which are Polynomial kernel, Linear Kernel and RBF. For this research, RBF kernel tailored as the kernel function of SVM as shown in formula (12).

$$k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2) \quad (12)$$

In this case γ is proportional inverse to the width of the kernel. In common, SVM classifier using a default value of C and γ on the to solve the classification problems.

E. Convolutional Neural Network (CNN) Classification

CNN known with its tremendous performance in image processing in and recognition[18]. As it nature as a part of multi-layer perceptron classifier, training stage is use to construct the weights and biases of the involved convolution kernels. Those weight and biases are defined through forward propagation using the training dataset, where the output O can be obtained. CNN can define the error E by comparing the output O along with the labels y. Given the sample dataset number is denoted by N, and the sample types number are c. The error E can be calculated according to the Equation (13).

$$E^N = \frac{1}{2} \|y^N - O^N\|^2 = \frac{1}{2} \sum_{n=1}^N \sum_{k=1}^c (y_k^n - O_k^n)^2 \quad (13)$$

As a final step, CNN decide the iteration convergence according to the constant set in value of E. The training stage is

over when the converge in certain value. If the result of iteration did not equal to convergence value, the result of iteration are set up in the residual δ . Related to activation function f , the residual function can be defined from the Eq.(14). In this case, sigmoid function is tailored as the activation function.

$$\delta^L = \frac{\delta}{\delta z_n} \frac{1}{2} \|y^n - O^n\|^2 = -(y^N - O^N) \cdot f'(Z^L) \quad (14)$$

The residues produced by output layer then propagated into front layer to re-calculate the residual in every layer (Eq. (15)), δ^l represent the residual result of the l^{th} layer.

$$\begin{aligned} \delta^l &= ((W^l - \alpha \frac{\partial}{\partial W^{(l)}} E(\omega, b) = W^{(l)} - \alpha \cdot \delta^{(l+1)} (a^{(l)})^T \\ b_i^{(l)} &= b_i^l - \alpha \cdot \frac{\partial}{\partial b_i^{(l)}} E(\omega, b) = b_i^l - \alpha \cdot \delta^{(l+1)} \end{aligned} \quad (15)$$

The iteration on CNN continuously execute on the way to found optimal result as set-up in the threshold constant.

III. RESEARCH METHODOLOGY

The steps of each classification process are shown in Fig. 3 for SVM-based and fig. 4 for CNN-based. Each classifier have their owned characteristic. Hence, the SVM has a slightly different steps which is equipped by feature extraction.

A. SVM-based Classifier

As shown in Fig.3, the classification started by data acquisition which is reading the dicom multiple slices medical images in certain location. Then those images through the grey scaling in data pre-processing stage. This step needed to simplify the computation. In the feature extraction step, Hu moment invariant is used to find the edge of lung. This mathematical morphology algorithm shows outstanding performance in handling shape-related processing and operations[14].

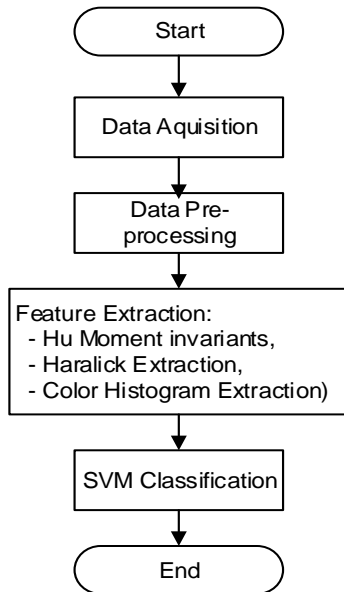


Fig. 3. Algorithm steps for SVM Classification

Meanwhile, Haralick is tailored to extract the color texture. The color texture classification needed in regrouping images which have similar texture. Then color histogram is used to measure histogram images. In this research the histogram consists of eight bins, where bin 1 consists of value 0-7, bin 2 consists of value 8-15 and so on until bin 8. The last step is SVM classification. The same steps are done during the training and testing steps.

B. CNN-based Classifier

This research employs three layers of convolutional network for its classification. CNN has more simple steps in general, but it has more complex computation. A number of iterations needed to find an optimal weight of each neuron. The algorithm started by data acquisition by reading dicom images. Those images are need to convert to grey scale mode for better computation time. Then the algorithm is trained by labeled data to find the optimal weights. After the training stage, the same steps are conduct for the testing stage. Fig. 4 shows the algorithm steps.

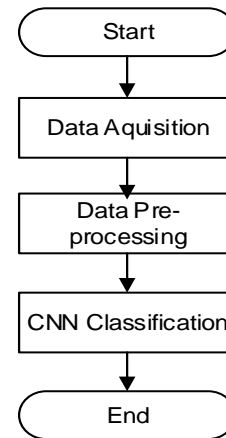


Fig. 4. Algorithm steps for CNN Classifier

IV. EXPERIMENT RESULT

Dataset involved in this research are computed tomography (CT) scans repository from Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI). Those dataset prepared for research purpose of medical imaging research community. This repository consist of 1700 image slices that are involve in during training and testing stage of this research. 50 patient dataset are use during training and 20 patient use during testing stages.

TABLE I. COMPARATION RESULT

Classifier	Accuracy (%)	Computation Time (second)
SVM	53	20
CNN	89	58

The machine involved in this research has an intel core 7th generation with 2.7 GHz clock speed, 8 Gigabyte of internal memory, and equipped with parallel Graphical Processing unit.

The results of experiment are shown in Table 1. The computation time results are the summary of training and testing stage for each classifier.

SVM-based classifier has a better speed with 20 seconds compared to CNN that spend 58 seconds for both training and testing sessions. This SVM result still have a better computation time although it has more step which is feature extraction. It can be concluded that feature extraction is not shortage of this classifier and become a common behavioral part needed by this classifier. In this research, CNN does not have any additional pre-classification steps such as feature extraction, but it still time consuming. Most of the time of this three-layer convolution algorithm is spent in training stage. It is cause by multiple iteration of image rotation and dilation to make sure that the algorithm is not trapped in over or under fitting condition. It is mean that this algorithm needs a multiple computation in the training stage that make it more time consuming.

CNN show it superiority in accuracy which has 89% compared to only 53% by SVM. The SVM classifier's accuracy can be improved by adjusting some variable in feature extraction stages. For example, this research only used 8 bins in color histogram extraction. The accuracy can be improved by adding more bins to store all of the color values, so it become more precise. This adjustment will have an impact to computation time in both of training and testing stages.

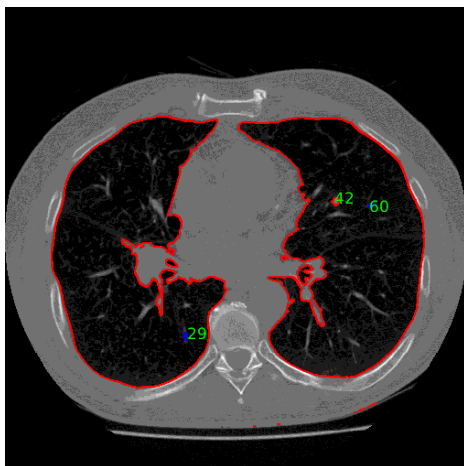


Fig. 5. Sample of Classification results of SVM

The example of lung nodule detection by SVM is shown by Fig5. The red line is an edge for the lung area. This area are detecting in the feature extraction of hu moment invariant step. The blue round-tiny shape are the nodule that detected by this classifier.

V. CONCLUSION

Both of SVM and CNN-based classifier have already implemented to detect lung nodule in CT scan images. The result shows that SVM classifier has a better computing speed although it has more step its algorithm. Meanwhile CNN has a better accuracy compare to SVM. SVM is very depending on the feature extraction steps. Adjusting some variable in that

step can improve the accuracy of lung detection in SVM, although it will have an impact to computation time.

ACKNOWLEDGMENT

This project is funded by Ministry of Research, Technology and Higher Education, under the prime fundamental research schema. The ground truth identification was assisted by Radiologists from Tabanan General Hospital and Airlangga University Hospital.

REFERENCES

- [1] N. Mesanovic, M. Grgic, H. Huseinagic, M. Males, E. Skejic, and M. Smajlovic, "Automatic CT Image Segmentation of the Lungs with Region Growing Algorithm," *Int. J. Eng. Adv. Technol.*, vol. 4, no. 5, pp. 395–400, June 2015.
- [2] H. M. Orozco, S. Member, O. Osiris, V. Villegas, S. Member, and H. De Jes, "Lung Nodule Classification in CT Thorax Images using Support Vector Machines," [12th Mexican International Conference on Artificial Intelligence, 2013], pp. 277–282, November 2013
- [3] P. Bhuvanewari and A. B. Therese, "Detection of Cancer in Lung With K-NN Classification Using Genetic Algorithm," *Procedia Mater. Sci.*, vol. 10, pp. 433–440, 2015.
- [4] F. V. Farahani, A. Ahmadi, and M. H. F. Zarandi, "Lung Nodule Diagnosis from CT Images Based on Ensemble Learning," 2015.
- [5] K. Mya, M. Tun, and A. S. Khaing, "Feature Extraction and Classification of Lung," *Int. J. Eng. Res. Technol.*, vol. 3, no. 3, pp. 2204–2210, September 2014.
- [6] C. Bhuvanewari, P. Aruna, and D. Loganathan, "Classification of Lung Diseases by Image Processing Techniques Using Computed Tomography Images," *Int. J. Adv. Comput. Res.*, vol. 4, no. 1, pp. 2277–7970, 2014.
- [7] J. Liu, J. Cai, K. Chellamuthu, M. Bagheri, L. Lu, and R. M. Summers, "Cascaded coarse-to-fine convolutional neural networks for pericardial effusion localization and segmentation on CT scans," [Int. Symp. Biomed. Imaging], vol. 2018–April, no. Isbi, pp. 1092–1095, 2018.
- [8] X. Liu, F. Hou, H. Qin, and A. Hao, "Multi-view multi-scale CNNs for lung nodule type classification from CT images," *Pattern Recognit.*, vol. 77, pp. 262–275, May 2018.
- [9] J. Guo, H. Du, J. Zhu, T. Yan, and B. Qiu, "Relative location prediction in CT scan images using convolutional neural networks," *Comput. Methods Programs Biomed.*, vol. 160, pp. 43–49, July 2018.
- [10] M. Larsson, Y. Zhang, and F. Kahl, "Robust abdominal organ segmentation using regional convolutional neural networks," *Appl. Soft Comput. J.*, vol. 70, pp. 465–471, 2017.
- [11] H. R. Tizhoosh and F. Khalvati, "Medical Image Classification via SVM using LBP Features from Saliency-Based Folded Data," pp. 0–4, 2015.
- [12] P. Eskandarian and J. Bagherzadeh, "Computer-Aided Detection of Pulmonary Nodules based on SVM in Thoracic CT Images," [7th International Conference on Information and Knowledge Technology], pp. 0–5, Desember 2015.
- [13] M. Keshani, Z. Azimifar, F. Tajeripour, and R. Boostani, "Lung nodule segmentation and recognition using SVM classifier and active contour modeling: A complete intelligent system," *Journal Computer Biology Medic.*, vol. 43, pp. 287–300, January 2013.
- [14] B. Vanajakshi, "Classification Of Boundary And Region Shapes Using Hu-Moment Invariants," *Indian J. Comput. Sci. Eng.*, vol. 3, no. 2, pp. 386–393, April 2012.
- [15] A. Porebski, N. Vandenbroucke, and L. Macaire, "Neighborhood and Haralick feature extraction for color texture," *Sci. Technol.*, no. 2, pp. 316–321, June 2008.
- [16] F. Alamdar and M. R. Keyvanpour, "A new color feature extraction method based on quad histogram," *Procedia Environ. Sci.*, vol. 10, no. PART A, pp. 777–783, 2011.
- [17] P. Chen, L. Yuan, Y. He, and S. Luo, "An improved SVM classifier

based on double application in analogue circuit diagnosis Reference : To appear in : Neurocomputing,” Neurocomputing, October 2015.

improved Convolutional Neural Network,” J. Vis. Commun. *Image Represent.*, vol. 55, no. May, pp. 640–647, August 2018.

[18] J. Zhang, K. Shao, and X. Luo, “Small sample image recognition using