

# Efektivitas Penggunaan Seleksi Ciri CFS pada Klasifikasi Ciri Bentuk Nodul Kanker Payudara dengan Citra Ultrasonografi

H.K.N.Yusufiyah<sup>1\*</sup> dan J.P.G.N.Rochman<sup>2</sup>

<sup>1</sup>*Biomedical Engineering Group Research, Surabaya, Indonesia*

<sup>2</sup>*Departemen Teknik Geofisika, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia*

\* Email: [hestikhuzaimah@gmail.com](mailto:hestikhuzaimah@gmail.com)

*Received: October 25<sup>th</sup>, 2020. Accepted: February 17<sup>th</sup>, 2021. Published: February 28<sup>th</sup>, 2021*

## Abstrak

Penggunaan ciri bentuk nodul merupakan salah satu parameter yang digunakan pada penentuan keganasan kanker payudara. Perhitungan secara matematis digunakan sebagai *second decision* untuk menguatkan ahli radiolog dalam menentukan keganasan kanker payudara menggunakan citra ultrasonografi (USG). Metode yang digunakan pada penelitian ini adalah menapis citra USG yang terdapat derau *speckle*, kemudian dilanjutkan proses segmentasi, penggalan ciri bentuk, seleksi ciri bentuk, serta klasifikasi. Proses seleksi ciri dengan *Correlated based Feature Selection* (CFS) digunakan untuk menyeleksi ciri bentuk yang dominan pada citra. Hasil klasifikasi yang didapatkan menunjukkan hasil seleksi ciri menggunakan CFS dapat meningkatkan hasil akurasi, sensitivitas serta spesifisitas citra, sehingga mampu membedakan ciri bentuk nodul kanker dengan lebih baik.

## Abstract

The implementation of nodule shape characteristics is one of the parameters used in determining breast cancer malignancy. Mathematical calculations are used as a second decision to strengthen radiologists in determining breast cancer malignancy using ultrasound images (USG). The method used in this research is to filter ultrasound images that contain speckle noise, then continue the segmentation process, extracting shape features, selecting shape features, and classifying them. The feature selection process using *Correlated*

based Feature Selection (CFS) is used to select the dominant shape features in the image. The classification results obtained show that the results of feature selection using CFS can improve the results of image accuracy, sensitivity and specificity, so as to be able to better distinguish the characteristic shape of the cancer nodule. ©2021PERJ

**Keywords:** Ultrasound; Breast cancer; CFS; Nodules.

## PENDAHULUAN

Berdasarkan data yang didapatkan dari WHO pada tahun 2018, angka kematian dikarenakan kanker payudara terdapat sekitar 11,6% (WHO, 2018). Provinsi Jawa Tengah menjadi salah satu provinsi dengan penderita kanker payudara terbanyak. Hal tersebut sesuai dengan data yang diperoleh dari laporan Kemenkes RI (Kemenkes RI, 2015).

Pemeriksaan menggunakan ultrasonografi (USG) menjadi salah satu cara deteksi primer untuk kanker payudara, selain menggunakan pemeriksaan sendiri dan mamografi (Yusufiyah *et al.*, 2015). Akan tetapi, saat ini terdapat kekurangan dalam pemeriksaan menggunakan USG, yaitu akurasi yang masih rendah (Yusufiyah *et al.*, 2015). Hal tersebut dikarenakan masih terdapat derau *speckle* pada citra USG. Pemeriksaan menggunakan USG juga harus diverifikasi terlebih dahulu oleh ahli radiolog untuk memastikan keakurasian pembacaan hasil citra USG (Kuo *et al.*, 2002; Chen *et al.*, 2003; Huang *et al.*, 2008). Oleh

karena itu, dibutuhkan suatu metode yang dapat membantu ahli radiolog dalam memverifikasi hasil USG secara matematis (Chen *et al.*, 2003). CAD (*Computer Aided Decission*) merupakan sistem yang mampu menjadi pertimbangan ke dua (*second decision*) yang sudah dikembangkan oleh ahli (Huang *et al.*, 2008; Wu and Moon, 2008; Yang, Chang, Huang, Chou, & Li, 2007). Akan tetapi, salah satu kekurangan sistem ini adalah belum adanya kemampuan klasifikasi secara independen terhadap salah satu ciri nodul kanker payudara, terutama ciri bentuk. Ciri bentuk menjadi penting untuk diketahui karena menjadi salah satu parameter secara visual untuk penentuan keganasan kanker payudara. Oleh karena itu, dibutuhkan suatu penelitian yang mampu mengklasifikasi bentuk nodul, agar mampu membantu dokter atau radiolog sebagai *second decision* dalam penentuan keganasan kanker payudara.

Penelitian sebelumnya sudah mampu mengklasifikasikan ciri bentuk nodul kanker payudara, dengan menggunakan karakteristik 35 ciri (H. A. Nugroho, Yusufiyah,

Adji, & Nugroho, 2015). Kekurangan penelitian ini belum mampu menyeleksi karakteristik ciri yang kurang dominan, serta belum mampu membandingkan hasil klasifikasi antar metode pengkelasan. Oleh karena itu, pada penelitian ini digunakan seleksi ciri untuk mereduksi ciri citra yang kurang dominan dengan membandingkan beberapa metode pengkelasan, sehingga mampu menghasilkan klasifikasi yang lebih baik dengan ciri yang lebih sedikit.

## METODE

Proses mereduksi derau *speckle* menggunakan metode adaptive median filter dan metode tapis *Speckle Reduction Bilateral Filtering* (SRBF). Kedua metode tapis ini sudah terbukti dapat mengurangi derau *speckle* (Yusufiyah, Nugroho, Adji, & Nugroho, 2015; Husna, Nugroho, & Soesanti, 2015; S. Wu, Zhu, & D, 2013; A. Nugroho & Nugroho, 2015; Balocco, Gatta, Pujol, Mauri, & Radeva, 2010). Tahapan dalam memisahkan latar belakang citra dan latar depan citra (proses segmentasi) menggunakan metode *active contour*. Metode ini juga diklaim mampu mensegmen citra secara baik (Yusufiyah et al., 2015; Husna et al., 2015; A. Nugroho & Nugroho, 2015).

Proses penggalan ciri bentuk menggunakan beberapa metode, yaitu momen invarian, momen zernike, parameter kebulatan, dan parameter kerampingan (Yusufiyah

et al., 2015). Metode tersebut digunakan kembali dalam penelitian ini karena terbukti mampu memberikan hasil klasifikasi yang baik dalam penelitian tersebut.

Proses seleksi ciri menggunakan metode *Correlation Based Feature Selection* (CFS) yang sudah terdapat pada *machine learning* WEKA. Prinsip kerja dari algoritme ini adalah dengan menemukan subset ciri dengan tidak menurunkan nilai dimensi dari dataset yang sudah ada. Hal ini diklaim mampu meningkatkan hasil akurasi pada klasifikasi data.

Metode CFS menggunakan matriks korelasi agar mampu mendapatkan nilai korelasi ciri yang tinggi pada kelas tersebut, tetapi ciri tersebut juga tidak berkorelasi dengan ciri lainnya. Hal ini yang disebut dengan koefisien korelasi Pearson. Algoritme ini dapat dilihat seperti pada Persamaan (1).

$$Ms = \frac{k\bar{r}_{cf}}{\sqrt{k + k(k-1)\bar{r}_{ff}}} \quad (1)$$

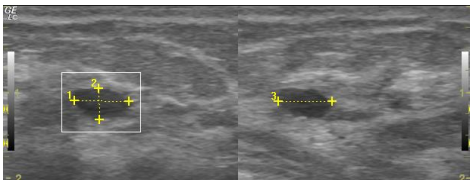
Pada algoritme Persamaan 1,  $M_s$  adalah "Merit" yang artinya berapa banyak fitur  $S$  (korelasi antara jumlah subset ciri dan variabel kelas), sedangkan  $k$  adalah jumlah ciri.  $\bar{r}_{cf}$  adalah rata-rata korelasi antara masing-masing ciri dengan kelasnya, serta  $\bar{r}_{ff}$  adalah rata-rata korelasi yang berpasangan antara dua ciri.

## HASIL DAN PEMBAHASAN

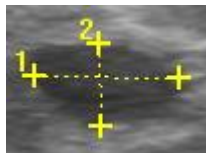
### Hasil Pemrosesan Citra

Pada tahap pertama, citra Ultrasonografi (USG) yang didapatkan diambil di daerah *Region of Interest* (ROI) nodul, seperti pada Gambar 1. Pengambilan daerah nodul ini sudah berdasarkan penanda yang dibuat oleh dokter radiologi.

Hasil pemotongan daerah ROI dijadikan citra yang diproses pada tahap selanjutnya, yaitu tahap penapisan citra, segmentasi, penggalan ciri, klasifikasi nodul, serta seleksi ciri. Citra pada Gambar 2 adalah contoh hasil ROI nodul. Hasil ROI nodul ini masih didapatkan secara manual, setelah konfirmasi penanda nodul oleh dokter.



Gambar 1. Citra Ultrasonografi (USG)

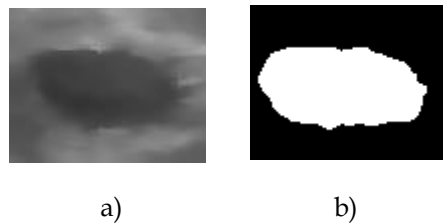


Gambar 2. Hasil ROI Citra Ultrasonografi (USG)

### Hasil Penapisan dan Segmentasi

Pada tahap ini, proses penapisan dilakukan dengan dua metode, yaitu metode *adaptive median filter* dan metode *speckle reduction bilateral filtering* (SRBF). Hasil citra tapis ini

kemudian disegmentasi dengan metode *active contour*. Metode ini dapat menghasilkan citra yang jelas pada daerah nodul yang akan diidentifikasi. Proses penapisan dan segmentasi dapat dilihat seperti pada Gambar 3. Hasil yang didapatkan seperti pada penelitian sebelumnya, citra mampu tersegmen dengan baik.



Gambar 3. a) Citra Hasil Penapisan, dan b) Citra Hasil Segmentasi

### Hasil Penggalan Ciri Bentuk Nodul Citra dan Seleksi Ciri

Pada tahap ini, citra yang sudah tersegmen, dilakukan penggalan ciri bentuk nodul dengan empat metode, yaitu metode momen invarian, momen zernike, parameter kebulatan, dan parameter kerampingan. Penggalan ciri tersebut dimaksudkan untuk mendapatkan nilai matematis pada ciri bentuk nodul, agar dapat digunakan untuk proses klasifikasi citra dan juga proses seleksi ciri citra. Pada proses ini, didapatkan 37 ciri bentuk nodul untuk diproses dengan menggunakan *machine learning WEKA*. 37 ciri tersebut terdiri atas tujuh ciri momen invarian, 28 ciri momen zernike, satu ciri parameter kebulatan, dan

satu ciri parameter kerampingan citra nodul.

Setelah didapatkan hasil penggalan ciri, kemudian dilakukan uji klasifikasi. Pada hasil klasifikasi ini dapat dilihat hasil akurasi, sensitivitas, serta spesifisitas ciri bentuk nodul tersebut seperti pada Tabel 1. Akan tetapi, hasil klasifikasi yang didapatkan kurang optimal, sehingga dilakukan uji seleksi ciri, agar mendapatkan ciri yang dominan. Seleksi ciri ini menggunakan metode *Correlation based Feature Selection* (CFS). Hasil

seleksi ciri tersebut didapatkan sebelas ciri yang dominan, yang terdiri atas satu ciri momen invarian, delapan ciri momen zernike, satu ciri parameter kebulatan, dan satu ciri parameter kerampingan.

Hasil seleksi ciri menggunakan metode CFS juga dapat dilihat pada Tabel 1. Pada uji klasifikasi, baik pada saat ciri sebelum diseleksi dan ciri sudah diseleksi, dilakukan pengujian dengan menggunakan beberapa metode klasifikasi, yaitu metode *classifier* MLP SMO, Naive bayes, dan J48.

**Tabel 1.** Hasil Uji Klasifikasi (Oval vs Bulat vs Ireguler) dengan Berbagai Jenis *Classifier*

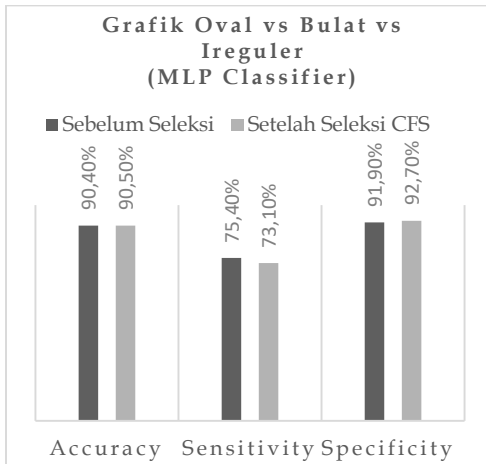
Jumlah Data	Jumlah Ciri	Akurasi (%)	Sensitivitas (%)	Spesifisitas (%)	Jenis <i>Classifier</i>
105	37	90,4	75,4	91,9	MLP
	11 (Setelah Seleksi CFS)	90,5	73,1	92,7	
	37	91,1	73,8	92,7	SMO
	11 (Setelah Seleksi CFS)	94,3	78,3	94,6	
	37	81,6	74,7	77,3	Naive Bayes
	11 (Setelah Seleksi CFS)	86,7	80,4	91,4	
	37	82,9	59,9	86,5	J48
	11 (Setelah Seleksi CFS)	85,4	60,4	87,4	

Pada Tabel 2 terlihat hasil klasifikasi ciri bentuk nodul dengan tiga kelas, yaitu bentuk bulat, oval, dan ireguler. Melalui hasil uji klasifikasi tersebut didapatkan nilai akurasi, sensitivitas, dan spesifisitas tertinggi terlihat pada penggunaan metode pengkelasan SMO, kemudian metode pengkelasan

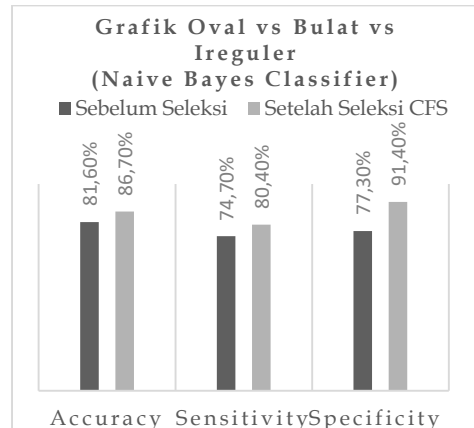
MLP. Pengkelasan menggunakan metode MLP tidak menjadi *classifier* terbaik dalam uji klasifikasi pada kasus dan data ini. Hal ini berbeda dengan hasil yang didapatkan pada penelitian sebelumnya yang menyatakan bahwa MLP merupakan *classifier* terbaik (Yusufiyah et al., 2015). Secara

analisis, hal ini dapat terjadi dikarenakan data-data yang digunakan dalam uji klasifikasi, sehingga hasil klasifikasi menggunakan metode MLP tidak menjadi *classifier* terbaik. Akan

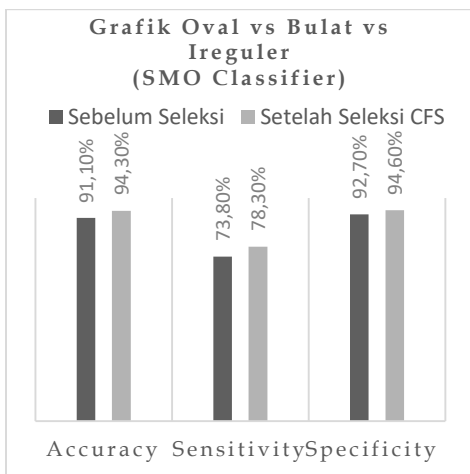
tetapi, hasil uji klasifikasi antara *classifier* MLP dan SMO juga tidak terlalu berbeda jauh, sehingga *classifier* MLP juga dapat digunakan untuk klasifikasi pada kasus-kasus yang sama dengan penelitian ini.



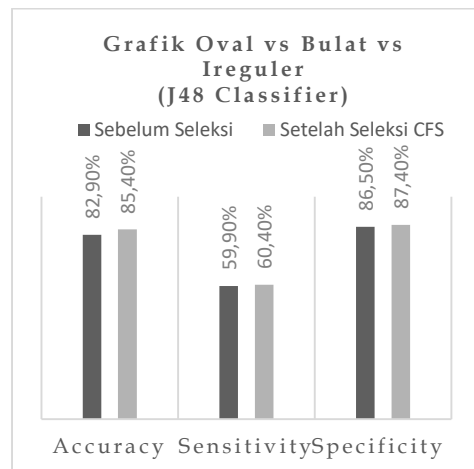
**Gambar 3.** Klasifikasi Oval vs Bulat vs Ireguler menggunakan MLP Classifier



**Gambar 5.** Klasifikasi Oval vs Bulat vs Ireguler menggunakan Naive Bayes Classifier



**Gambar 4.** Klasifikasi Oval vs Bulat vs Ireguler menggunakan SMO Classifier



**Gambar 6.** Klasifikasi Oval vs Bulat vs Ireguler menggunakan J48 Classifier

Gambar 3, 4, 5, dan 6, menunjukkan bahwa hasil akurasi, sensitivitas, dan spesifisitas semakin meningkat dengan menggunakan

sebelas ciri bentuk nodul. Ciri tersebut adalah ciri yang didapatkan setelah dilakukan seleksi ciri menggunakan CFS. Oleh karena itu,

dapat dikatakan bahwa metode seleksi ciri dengan CFS mampu meningkatkan nilai akurasi, sensitivitas, dan juga spesifisitas pada masing-masing metode *classifier*.

## KESIMPULAN

Hasil seleksi ciri dengan menggunakan metode *Correlation Based Feature Selection* (CFS) terbukti efektif dalam membantu peningkatan hasil klasifikasi bentuk nodul citra. Hal tersebut terlihat pada hasil klasifikasi (akurasi, sensitivitas, serta spesifisitas) yang semakin meningkat setelah dilakukan seleksi ciri bentuk nodul dengan CFS. Selain itu, klasifikasi menggunakan metode klasifikasi SMO dan MLP terbukti dapat memberikan hasil yang lebih tinggi dibanding metode klasifikasi yang lain.

## DAFTAR PUSTAKA

- Balocco, S., Gatta, C., Pujol, O., Mauri, J., & Radeva, P. 2010. SRBF: Speckle reducing bilateral filtering. *Ultrasound in Medicine & Biology*, 36(8), 1353-63. <https://doi.org/10.1016/j.ultrasmedbio.2010.05.007>
- Chen, C., Chou, Y., Han, K., Tiu, C., Chiou, H., & Chiou, S. 2003. Breast Lesions on Sonograms : Computer-aided Diagnosis with Nearly Setting- Independent Features and Artificial Neural Networks 1. *Radiology*, 226(7).
- Huang, Y.-L., Chen, D.-R., Jiang, Y.-R., Kuo, S.-J., Wu, H.-K., & Moon, W. K. 2008. Computer-aided diagnosis using morphological features for classifying breast lesions on ultrasound. *Ultrasound Obstet Gynecol*, 32(4), 565-572. <https://doi.org/10.1002/uog.5205>
- Husna, D. A., Nugroho, H. A., & Soesanti, I. 2015. Performance Analysis of Edge and Detail Preserved Speckle Noise Reduction Filters for Breast Ultrasound Images. In *The 2nd International Conference on Information Technology, Computer, and electrical Engineering* (pp. 79-83). Semarang: Departement of Computer Engineering, Diponegoro University.
- Kemenkes RI, I. 2015. *Situasi penyakit kanker*. Jakarta Selatan.
- Kuo, W., Chang, R., Moon, W. K., Lee, C. C., & Chen, D. 2002. Computer-aided Diagnosis of Breast Tumors with Different US Systems 1. *Academic Radiology*, 9(1), 793-799.
- Nugroho, A., & Nugroho, H. A. 2015. Active Contour Bilateral Filtering for Breast Lesions Segmentation on Ultrasound Images. In *2015 International Conference on Science in Information Technology (ICSITech)*. Yogyakarta. <https://doi.org/978-1-4799-8385-8>
- Nugroho, H. A., Yusufiyah, H. K. N.,

- Adji, T. B., & Nugroho, A. 2015. Zernike Moment Feature Extraction for Classifying Lesion's Shape of Breast Ultrasound Image.
- WHO. 2018. *Latest global cancer data : Cancer burden rises to 18 . 1 million new cases and 9 . 6 million cancer deaths in 2018.* france. Retrieved from <https://www.uicc.org/iarc-release-latest-world-cancer-statistics>
- Wu, S., Zhu, Q., & D, Y. X. P. 2013. Evaluation of Various Speckle Reduction Filters on Medical Ultrasound Images. In *35th Annual International Conference of the IEEE EMBS* (pp. 1148-1151). Osaka.
- Wu, W.-J., & Moon, W. K. 2008. Ultrasound breast tumor image computer-aided diagnosis with texture and morphological features. *Academic Radiology*, 15(7), 873-80. <https://doi.org/10.1016/j.acra.2008.01.010>
- Yang, H., Chang, C., Huang, S., Chou, Y., & Li, P. 2007. Breast Ultrasound Computer-Aided Diagnosis Using Both Acoustic and Image Features. *IEEE Ultrasonics Symposium*.
- Yusufiyah, H. K. N., Nugroho, H. A., Adji, T. B., & Nugroho, A. 2015. Feature Extraction for Classifying Lesion ' s Shape of Breast Ultrasound Images. In *The 2nd International Conference on Information Technology, Computer, and electrical Engineering* (pp. 105-109). Departement of Computer Engineering, Diponegoro University.