

## **Sistem Pengelompokan Jenis Kaleng Bekas Berdasarkan Model Warna CMYK**

### *Grouping System of the Used Cans Type Based on CMYK Color Model*

**Yulia Resti**<sup>1\*)</sup>, F. Burlian<sup>2</sup>, Fusito Fusito<sup>2</sup>, I. Yani<sup>2</sup>

<sup>1</sup>Jurusan Matematika, Fakultas MIPA, Universitas Sriwijaya, Sumatera Selatan 30862

<sup>2</sup>Jurusan Teknik Mesin, Fakultas Teknik, Universitas Sriwijaya, Sumatera Selatan 30862

<sup>\*)</sup>Penulis untuk korespondensi: yulia\_resti@mipa.unsri.ac.id

**Sitasi:** Resti Y, Burlian F, Fusito F, Yani I. 2019. Grouping system of the used cans type based on CMYK color model. *In: Herlinda S et al. (Eds.), Prosiding Seminar Nasional Lahan Suboptimal 2019, Palembang 4-5 September 2019.* pp. 338-344. Palembang: Unsri Press.

### **ABSTRACT**

The automatic sorting process in the recycling industry requires a grouping system. This research aims to build a grouping system of used cans into three types; aluminum, tin plate and aerosols. The grouping is done based on the Cyan, Magenta, Yellow, and BLacK (CMYK) color models of the image of the used cans which are captured using a digital camera when placed on a conveyor belt. The grouping method used is a statistical method based on the Bayes theorem and the assumption of independence on the input variables. The results obtained show that the Naïve Bayes method is able to group used cans by type with an accuracy rate greater than 82%.

---

Keywords: accuracy rate, digital images, naïve bayes

### **ABSTRAK**

Proses sortir otomatis dalam industri daur ulang memerlukan sistem pengelompokan. Penelitian ini bertujuan untuk membangun sistem pengelompokan kaleng bekas ke dalam tiga kelompok; aluminium, tin plate, dan aerosol. Pengelompokan dilakukan berdasarkan model warna Cyan, Magenta, Yellow, dan black (CMYK) dari citra kaleng bekas yang dicaptur menggunakan kamera digital pada saat diletakkan di atas belt conveyor. Metode pengelompokan yang digunakan adalah metode statistik berbasis teorema Bayes dan asumsi independensi pada variabel inputnya. Hasil penelitian yang diperoleh menunjukkan bahwa metode Naïve Bayes mampu mengelompokkan kaleng bekas menurut jenisnya dengan tingkat akurasi lebih besar daripada 82%..

---

Kata kunci: citra digital, naïve bayes, tingkat akurasi

### **PENDAHULUAN**

Citra digital, secara sederhana, dapat dipahami sebagai gambar dua dimensi yang dapat ditampilkan pada monitor. Suatu citra digital tersusun dari titik-titik yang disebut dengan piksel. Setiap piksel memiliki warna dan warna tersebut disusun dari tiga warna primer, yaitu merah, hijau dan biru. Model warna yang menggunakan ketiga warna primer ini dikenal sebagai model warna RGB. Menurut Dubey *et al.*, (2013) teknik citra digital merupakan salah satu cara untuk mengakses informasi pada sebuah gambar. Teknik citra digital ini dapat digunakan untuk mengelompokkan suatu objek tertentu. Sistem pengelompokan objek berdasarkan citra digital dapat dilakukan menggunakan model warna RGB maupun model warna Cyan (hijau kebiruan), Magenta (merah keunguan),

Yellow (kuning), dan black (hitam) yang lazim disingkat CMYK. Apabila dibandingkan dengan model warna RGB, model warna CMYK memiliki spektrum warna yang lebih luas.

Sistem pengelompokan berdasarkan citra digital telah banyak diterapkan pada berbagai industri, diantaranya pada industri daur ulang yang menerapkan sistem otomatisasi. Paper ini memaparkan sistem pengelompokan kaleng bekas berdasarkan citra digital ke dalam 3 kelompok yaitu tin plate, aluminium, dan aerosol dimana model warna yang digunakan adalah model warna CMYK dimana paper-paper sebelumnya menggunakan model warna RGB (Yani *et al.*, 2009; Yani *et al.*, 2013; Resti, 2015; dan Resti *et al.*, 2017), sedangkan metode pengelompokan yang digunakan adalah metode Naïve Bayes. Metode berbasis teorema Bayes dengan asumsi independensi yang kuat ini (John & Langley, 1995; Friedman *et al.*, 1997 dalam Aziz *et al.*, 2016), seringkali mengungguli metode lain untuk dataset yang besar (Sequira *et al.*, 2017 dan Revathi & Jawahar, 2015). Penelitian ini bertujuan untuk membangun sistem pengelompokan kaleng bekas ke dalam tiga kelompok; aluminium, tin plate, dan aerosol.

## BAHAN DAN METODE

Data citra digital kaleng diperoleh dengan menempatkan kaleng-kaleng di atas belt conveyor dan mengambil gambarnya pada pose depan, belakang, dan samping dengan menggunakan sebuah kamera digital yang dihubungkan dengan komputer menggunakan bahasa pemrograman Matlab 2016a. Dari gambar-gambar tersebut diperoleh rata-rata nilai piksel model warna primer yang kemudian ditransformasi ke model warna CMYK.

Misalkan  $X_1, X_2, X_3$ , dan  $X_4$  adalah variable-variabel acak yang merepresentasikan nilai transformasi piksel ke warna Cyan, Magenta, Yellow dan black. Keempat variabel tersebut bebas bersyarat satu terhadap yang lain. Misalkan  $P(T_j)$  merupakan peluang prior dari tipe kaleng ke- $j$ , dan  $P(X_1, X_2, X_3, X_4 | T_j)$  merupakan fungsi likelihood dari tipe kaleng ke- $j$ . Berdasarkan teorema Bayes, peluang posterior dari tipe kaleng ke- $j$  bersyarat pada  $(X_1, X_2, X_3, X_4)$  dapat didefinisikan sebagai (John & Langley, 1995; Mitchell, 2005; dan Zaidi *et al.*, 2013).

$$P(T_j | X_1, X_2, X_3, X_4) = \frac{P(X_1, X_2, X_3, X_4 | T_j)P(T_j)}{P(X_1, X_2, X_3, X_4)} \quad (1)$$

Pada persamaan (1),  $P(X_1, X_2, X_3, X_4)$  bernilai sama untuk setiap jenis kaleng sehingga penyebut ini dapat diabaikan dan persamaan (1) menjadi,

$$P(T_j | X_1, X_2, X_3, X_4) = P(X_1, X_2, X_3, X_4 | T_j)P(T_j) \quad (2)$$

Asumsi naïve pada teorema Bayes mengakibatkan peluang prior dapat dinyatakan sebagai,

$$P(X_1, X_2, X_3, X_4 | T_j) = \prod_{k=1}^4 P(X_k | T_j) \quad (3)$$

akibatnya (2) dapat ditulis sebagai,

$$P(T_j | X_1, X_2, X_3, X_4) = \prod_{k=1}^4 P(X_k | T_j)P(T_j) \quad (4)$$

Misalkan  $X_k$  merupakan variable acak yang berdistribusi normal dengan rata-rata,  $\mu_k$ , dan simpangan baku,  $\sigma_k$ , dimana  $k = 1,2,3,4$ , untuk masing-masing tipe kaleng  $T_j$ , fungsi densitas peluangnya dapat dituliskan sebagai,

$$P(X_k|T_j) = \frac{1}{\sigma_k\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_k - \mu_k}{\sigma_k}\right)^2\right] \quad (5)$$

dimana parameter,  $\mu_k$ , dan,  $\sigma_k$  pada (5) dapat diperoleh menggunakan metode moment atau maksimum likelihood.

Sebuah kaleng dapat dikelompokkan ke dalam tipe kaleng  $T_j$ , jika peluang posterior pada  $T_j$  seperti yang dirumuskan pada (4) merupakan peluang terbesar dibandingkan dengan peluang posterior pada tipe kaleng yang lain, yaitu.

$$\arg \max_{T_j} P(T_j|X_1, X_2, X_3, X_4) \quad (6)$$

Sampel penelitian terdiri 29,6% kaleng tipe 1 (tin plate), 33,2% kaleng tipe 2, dan 37,2% kaleng tipe 3. Sistem pengelompokan dengan metode multivariate naïve Bayes dilakukan berdasarkan komposisi antara data latih dan data uji dengan rasio 10 : 90, 20 : 80, 30 : 70, 40 : 60, 50 : 50, 60 : 40, 70 : 30, 80 : 20, dan 90 : 10.

## HASIL

Penyajian hasil penelitian diberikan dalam 3 bagian: hasil pengelompokan data latih, hasil pengelompokan data uji, dan tingkat akurasi pengelompokan.

### Hasil Pengelompokan Data Latih

Tabel 1 – Tabel 10 menyajikan hasil pengelompokan untuk data latih untuk masing-masing komposisi data 10 % - 100 %.

Tabel 1. Hasil pengelompokan untuk 10 % data latih (%)

		Tipe Kaleng Asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	48.00	0	0
	2	0	32.00	0
	3	0	0	20.00

Tabel 2. Hasil pengelompokan untuk 20% data latih (%)

		Tipe Kaleng Asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	38.00	0	0
	2	0	34.00	0
	3	0	0	28.00

Tabel 3. Hasil pengelompokan untuk 30% data latih (%)

		Tipe Kaleng Asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	36.00	1.33	0
	2	0	30.67	0
	3	0	0	32.00

Tabel 4. Hasil pengelompokan untuk 40% data latih (%)

		Tipe Kaleng Asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	36.00	1.00	0
	2	0	30.67	0
	3	0	0	33.00

Tabel 5. Hasil pengelompokan untuk 50% data latih (%)

		Tipe Kaleng Asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	32.00	8.00	0
	2	8.00	31.2	0
	3	0	0	35.20

Tabel 6. Hasil pengelompokan untuk 60% data latih (%)

		Tipe Kaleng Asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	30.00	0	0
	2	1.33	31.33	0
	3	0	2.00	35.33

Tabel 7. Hasil pengelompokan untuk 70% data latih (%)

		Tipe Kaleng Asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	29.14	0.57	0
	2	2.86	30.86	1.71
	3	0	1.14	33.71

Tabel 8. Hasil pengelompokan untuk 80% data latih (%)

		Tipe Kaleng Asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	29.00	0.50	0
	2	2.00	31.00	1.50
	3	0	1.00	35.00

Tabel 9. Hasil pengelompokan untuk 90% data latih (%)

		Tipe kaleng asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	29.33	0	0
	2	2.22	32.44	1.78
	3	0	20.89	33.33

Tabel 10. Hasil pengelompokan untuk 100% data latih (%)

		Tipe kaleng asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	27.20	0	0
	2	2.40	32.00	2.40
	3	0	1.20	34.80

### Hasil Pengelompokan Data Uji

Hasil pengelompokan untuk data uji untuk masing-masing komposisi data dapat dilihat pada Tabel 11 – Tabel 19.

Tabel 11. Hasil pengelompokan untuk 90 % data uji (%)

		Tipe kaleng asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	19.56	0.44	0
	2	3.11	29.78	5.78
	3	4.89	3.11	33.33

Tabel 12. Hasil pengelompokan untuk 80% data uji (%)

		Tipe kaleng asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	19.00	0	0
	2	2.00	32.00	6.50
	3	6.50	1.00	33.00

Tabel 13. Hasil pengelompokan untuk 70% data uji (%)

		Tipe kaleng asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	23.43	0.57	0
	2	3.43	32.00	5.14
	3	0	1.14	34.29

Tabel 14. Hasil pengelompokan untuk 60% data uji (%)

		Tipe kaleng asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	20.67	0.67	0
	2	3.33	32.67	6.00
	3	1.33	21.33	34.00

Tabel 15. Hasil pengelompokan untuk 50% data uji (%)

		Tipe kaleng asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	23.20	0.80	0
	2	3.20	32.00	4.00
	3	0	1.60	35.20

Tabel 16. Hasil pengelompokan untuk 40% data uji (%)

		Tipe kaleng asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	24.00	0	1.00
	2	3.00	32.00	4.00
	3	0	1.00	35.00

Tabel 17. Hasil pengelompokan untuk 30% data uji (%)

		Tipe kaleng asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	22.67	1.33	0
	2	1.33	33.33	4.00
	3	0	0	37.33

Tabel 18. Hasil pengelompokan untuk 20% data uji (%)

		Tipe kaleng asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	22.00	2.00	0
	2	2.00	34.00	6.00
	3	0	0	34.00

Tabel 19. Hasil pengelompokan untuk 10% data uji (%)

		Tipe kaleng asal		
		1	2	3
Pengelompokan kaleng ke tipe	1	8.00	0	0
	2	4.00	32.00	8.00
	3	0	0	48.00

### **Tingkat Akurasi Hasil Pengelompokan**

Tingkat akurasi pengelompokan untuk semua komposisi data latih berbanding data uji diberikan pada Tabel 20.

Tabel 20. Tingkat akurasi data

Komposisi Data Latih : Data Uji	Tingkat Akurasi	
	Data Latih (%)	Data Uji (%)
10 : 90	100.00	82.67
20 : 80	100.00	84.00
30 : 70	98.67	89.71
40 : 60	99.00	87.33
50 : 50	98.40	90.40
60 : 40	96.67	91.00
70 : 30	93.71	93.33
80 : 20	95.00	90.00
90 : 10	95.11	88.00
100		94.00

## **PEMBAHASAN**

Hasil pengelompokan pada data latih dengan komposisi sebesar 10% dan 20% yang ditampilkan masing-masing pada tabel 1 dan tabel 2 menunjukkan bahwa dengan menggunakan metode naïve Bayes semua kaleng berhasil terkelompok ke jenisnya dengan benar, sedangkan pada komposisi yang lain yang disajikan pada tabel 3-tabel 10 masih terdapat kaleng yang terkelompok ke jenis lain. Kebanyakan kaleng yang berasal dari jenis ke-2 yang masih terkelompok ke jenis kaleng ke-1 atau dan ke jenis kaleng ke-3. Pada semua komposisi data uji yang disajikan pada tabel 11-tabel 19 masih terdapat kaleng yang terkelompok ke jenis lain.

Tabel 20 memperlihatkan bahwa pada data latih, tingkat akurasi pengelompokan tertinggi adalah 100% yang dimiliki oleh data latih dengan komposisi 10% dan 20%, sedangkan tingkat akurasi pengelompokan terendah adalah 93.71% yang dimiliki oleh data latih dengan komposisi 70%. Pada data uji, tingkat akurasi pengelompokan tertinggi adalah 93.33% yang dimiliki oleh data uji dengan komposisi 30%, sedangkan tingkat akurasi pengelompokan terendah adalah 82.67% yang dimiliki oleh data latih dengan komposisi 90%. Hanya data dengan komposisi 50:50, 60:40, dan 70:30 yang memiliki tingkat akurasi pengelompokan di atas 90% di kedua-dua data latih dan data uji. Dari Tabel 20 dapat disimpulkan bahwa baik pada data latih maupun data uji tidak dapat dikatakan bahwa semakin besar persentase data, tingkat akurasi pengelompokan dengan metode multivariat naïve Bayes semakin rendah begitu juga sebaliknya semakin kecil persentase data, tingkat akurasi semakin tinggi.

## **KESIMPULAN**

Paper ini telah membangun sistem pengelompokan kaleng bekas berdasarkan citra digital dan model warna CMYK dengan menggunakan metode naïve Bayes dimana pengelompokan dilakukan berdasarkan komposisi antara data latih dan data uji dengan

rasio 10:90, 20:80, 30:70, 40:60, 50:50, 60:70, 80:10, dan 90:10. Tingkat akurasi pengelompokan yang diperoleh dari semua komposisi tersebut menunjukkan bahwa metode Naïve Bayes berhasil mengelompokkan kaleng ke dalam 3 jenis dengan tingkat akurasi lebih besar daripada 82% baik pada data latih maupun pada data uji. Tingkat akurasi tertinggi sebesar 100% berhasil dilakukan pada 10% dan 20% dari data latih.

## UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Universitas Sriwijaya sebagai penyandang dana penelitian Unggulan Kompetitif TA 2019 dan pihak-pihak yang berjasa dalam pelaksanaan penelitian ini.

## DAFTAR PUSTAKA

- Aziz R, Verma CK, Srivastava N. 2016. A fuzzy based feature selection from independent component subspace for machine learning classification of microarray data. *Genomics Data. Elsevier*. 8: 4-15.
- Dubey SR, Dixit P, Singh N. 2013. Infected Fruit Part Detection using K-Means Clustering Segmentation Technique. *International Journal of Artificial Intelligence and Interactive Multimedia*. pp. 65-72.
- John GH, Langley P. 1995. Estimating Continuous Distribution in Bayesian Classifiers. *Proceeding of the Eleventh Conference on Uncertainty in Artificial Intelligence*. San Mateo: Morgan Kaufmann Publisher.
- Langley, P., Iba, Wayne., dan K. Thompson. 1992. An Analysis of Bayesian Classifier. USA: NASA Ames Research Center.
- MatLab. 2016. MatLab: Numerical Computing. Tutorials Point (I) Pvt. Ltd. [http://mayankagr.in/images/matlab\\_tutorial.pdf](http://mayankagr.in/images/matlab_tutorial.pdf) [Diakses tanggal 21 Agustus 2019].
- Mitchell TM. 2005. Machine learning: Generative And Discriminative Classifiers: Naive Bayes And Logistic Regression.
- Revathi K, Jawahar M. 2015. Efficiency Improvement in Classification Tasks using Naive Bayes Tree and Fuzzy Logic. *International Journal For Trends In Engineering & Technology*. 6(1): 167-170.
- Resti Y. 2015. Dependence in Classification of Aluminium Waste, *Journal of Physics: Conference Series*. 622. 12052.
- Resti Y, Yani I, Mohruni AS, Burlian F, Amran A. 2017. A Probability Approach in Cans Identification. *Matec Web of Conference*.
- Sequera MS, Guirnaldo SA, Permites, ID. 2017. Naïve bayes classifier and fuzzy logic system for computer-aided detection and classification of mammamographic abnormalities. *Journal of Theoretical and Applied Information Technology*. 95(2): 381-392.
- Yani I. *et al.* 2009. Detecting Object Using Combination of Sharpening and Edge Detection Method, *European Journal of Scientific Research*. 32(1):121-127.
- Yani I. *et al.* 2013. An Automatic Sorting System for Recycling Beverage Cans using the Eigenface Algorithm,” *The Third International Conference on Soft Computing Technology in Civil, Structural and Environmental Engineering*. 103(21).
- Zaidi NA, Cerquides J, Charman MJ, Webb GI. 2013. Alleviating Naive Bayes Attribute Independence Assumption by Attribute Weighting. *Journal of Machine Learning Research*. 14 :1947-1988.