

PENERAPAN METODE *K-MEANS CLUSTERING* DAN *SUPPORT VECTOR MACHINE (SVM)* DALAM IDENTIFIKASI API PADA CITRA WARNA DIGITAL

Amanda Febrianti¹, Arini, S.T., M.T.², Feri Fahrianto, M.Sc.³

Teknik Informatika, Fakultas Sains dan Teknologi
UIN Syarif Hidayatullah Jakarta

Jl. Ir H. Juanda No.95, Cemp. Putih, Kec. Ciputat, Kota Tangerang Selatan, Banten 15412
manda.febriantii@gmail.com¹, arini@uinjkt.ac.id²
, feri.fahrianto@uinjkt.ac.id³

Abstrak - Terdapat sejumlah bencana yang terjadi di beberapa waktu ini, mulai dari bencana alam maupun bencana yang disebabkan oleh manusia itu sendiri. Salah satu bencana yang sering terjadi adalah Kebakaran. Dari kebakaran pada daerah pemukiman maupun kebakaran hutan. Upaya yang dapat dilakukan dalam mencegah bencana dari tahap paling awal. Metode yang dapat digunakan dalam upaya pendeteksian adalah pendeteksian pada citra digital, karena karakteristik api yang langsung bisa di dilihat secara visual. Penelitian ini akan fokus pada salah satu karakteristik visual api, yaitu warna. Untuk mendeteksi warna api, diperlukan tiga tahap yaitu segmentasi fitur warna, ekstraksi data, dan klasifikasi data. Untuk segmentasi warna, penulis menggunakan ruang warna YCbCr dan RGB. Ekstraksi fitur kemudian akan dilakukan dengan menggunakan metode K-Means Clustering. Hasil ekstraksi akan diklasifikasikan dengan metode Support Vector Machine (SVM) untuk mengelompokkan citra yang mengandung api dari citra biasa. Untuk data latih dan data uji, penulis akan menggunakan citra hutan agar dapat melakukan perbandingan hasil dengan penelitian sebelumnya. Hasil penelitian menunjukkan bahwa penggabungan mode warna RGB dan YCbCr dalam mendeteksi api pada citra digital hutan dan bangunan memiliki tingkat keberhasilan >90%, dengan tingkat akurasi yang lebih tinggi pada citra digital yang diambil pada siang hari.

Kata Kunci: *Pengolahan Citra, RGB, YCbCr, K-Means Clustering, Support Vector Machine, Api Ignition, Api Growth*

PENDAHULUAN

Banyak bencana yang telah terjadi akhir-akhir ini, mulai dari bencana alam maupun bencana buatan manusia itu sendiri. Menurut Badan Nasional Penanggulangan Bencana, bencana merupakan peristiwa atau rangkaian peristiwa faktor alam dan faktor non-alam maupun dari manusia yang dapat mengancam dan mengganggu kehidupan kehidupan

masyarakat seperti kerusakan lingkungan, korban jiwa manusia, dampak psikologis dan kerugian harta benda [1]

Dalam [2], kebakaran merupakan proses yang cepat dari oksidasi, reaksi eksotermis dimana bagian dari energi yang dilepaskan menyokong proses tersebut.

Pada [3], dijelaskan bahwa kebakaran merupakan fenomena yang terjadi pada bahan jika temperatur kritis dicapai yang bereaksi secara kimia dengan oksigen sehingga menghasilkan nyala api, panas, karbon monoksida, asap, karbon dioksida, uap air atau efek lainnya.

Kebakaran dapat terjadi di pemukiman, tempat umum, hutan, perkotaan maupun di kawasan industri.

Menurut NFPA (*National Fire Protection Association*) Amerika, tahap kebakaran terbagi menjadi empat, yaitu:

- *Ignition*: Bermula dari gabungan bahan bakar, oksigen, dan panas. Pada tahap ini, api masih dapat terkendali
- *Growth*: Dengan api awal sebagai sumber panas dan ditambah dengan bahan bakar, api akan menyebar ke lebih banyak permukaan.
- *Fully Developed*: Api telah menyebar ke semua bahan yang dapat terbakar dan tingkat temperatur telah mencapai puncaknya. Hal ini menyebabkan kerusakan lebih. Tingkat oksigen dalam ruang lingkup api pun akan berkurang drastis.
- *Burnout*: Api telah membakar seluruh permukaan yang dapat terbakar, sehingga api semakin lama akan semakin padam.

Pencegahan kebakaran sejak dini merupakan salah satu upaya yang dapat dilakukan dalam mengurangi kerusakan yang diakibatkan oleh kebakaran yang telah menyebar. Pencegahan dapat dilakukan dengan mengenali api dalam tahap *Ignition* dan *Growth*. Salah satu upaya yang dapat dilakukan untuk mengenali api pada tahap awal

adalah pendeteksian pada citra digital, dikarenakan karakteristik api yang bisa langsung dilihat secara visual. Salah satu karakteristik api adalah warna, gerakan, bentuk, dan tingkah laku asap [4]. Pada penelitian ini, penulis akan fokus pada salah satu karakteristik api, yaitu warna.

Dalam segmentasi fitur warna, penulis menggunakan ruang warna $YCbCr$ dan RGB . Pemilihan $YCbCr$ color space dalam ekstraksi warna berdasarkan penelitian S Rudz [5] dimana dijelaskan mengenai $YCbCr$ color space, yang didefinisikan sebagai ruang warna yang telah digunakan secara luas dalam bidang digital video dan broadcast, sehingga memudahkan dalam mendapatkan informasi visual video [6]. Dalam ekstraksi citra digital, color space lainnya didapatkan dari transformasi linier dan non-linier dari RGB . Ruang warna $YCbCr$ unggul dalam hal perbedaan pencahayaan yang dibutuhkan dalam pendeteksian kebakaran [7]. Menurut penelitian Emmy Premal, $YCbCr$ dapat memisahkan luminance dari chrominance, yang membuatnya unggul dalam memisahkan aspek pencahayaan daripada ruang warna lain seperti RGB [10].

Setelah tahapan normalisasi warna citra, data citra diolah dengan menggunakan teknik data mining. Salah satu metode *data mining* adalah pengelompokan data dan dalam penelitian ini, penulis akan menggunakan algoritma *K-Means Clustering*. Menurut Jain A. K [8], *K-Means* adalah salah satu algoritma yang sering digunakan karena kesederhanaannya. Oleh karena itu, penulis menggunakan algoritma *K-Means Clustering* untuk melakukan ekstraksi fitur citra.

Setelah data fitur pada citra digital di ekstraksi, diperlukan metode klasifikasi untuk mengelompokkan citra yang mengandung api dengan citra biasa. Penulis menggunakan metode klasifikasi *Support Vector Machine (SVM)*, berdasarkan pada penelitian pada tahun 2015 [9], dimana peneliti melakukan perbandingan *Learning Vector Quantization (LVQ)* dan *Support Vector Machine (SVM)* untuk prediksi penyakit jantung koroner, membuktikan bahwa *SVM* lebih baik dibandingkan *LVQ*. Kedua metode memberikan performansi klasifikasi yang hampir sama namun terlihat bahwa jaringan *SVM* memberikan tingkat akurasi yang lebih tinggi yaitu sebesar 90.3% dibandingkan jaringan *LVQ* sebesar 68.1%.

Kemudian menurut penelitian Zuoning Wang, menggabungkan ruang warna $YCbCr$ dan algoritma *K-Means* menghasilkan segmentasi gambar yang lebih akurat. Dengan memisahkan saluran warna menggunakan algoritma *K-Means*, Nilai abu-abu nyata api pada saluran kromatisitas dianalisis, dan *threshold* nyata api hutan diatur untuk menilai gambar apakah mengandung unsur api [11].

Berdasarkan dari studi literature yang telah digunakan berikut fokus penelitian kami :

1. Melakukan skenario penelitian citra yang diambil pada siang hari dan malam hari sebagai tolak ukur.
2. Pendeteksian api dalam tahap *Ignition* dan *Growth*, seperti yang dijelaskan oleh NFPA (*National Fire Protection Association*).
3. Menggabungkan fitur warna merah api (*Red*) dari ruang warna RGB dan fitur cahaya api (*Y*) dari ruang warna $YCbCr$ untuk memaksimalkan akurasi dalam mendeteksi api.
4. Fitur yang diekstraksi akan di klasifikasikan menggunakan *Support Vector Machine (SVM)* sebagai metode klasifikasi citra.
5. Menggunakan fitur warna yang telah diekstraksi dari ruang warna RGB dan $YCbCr$ sebagai data latih dan data uji. Hasil ekstraksi ruang warna akan dinormalisasi dengan mencari nilai rata-rata (*mean*) dari table hasil ekstraksi fitur.

METODOLOGI

A. Problem Formulation

Tahap pertama dalam metode simulasi komputer ini yaitu merumuskan permasalahan yang akan dianalisa dan menentukan penggunaan simulasi atau tidak untuk mendapatkan solusi dari permasalahan yang telah dirumuskan. Setelah mengumpulkan data, penulis merumuskan masalah bagaimana mengidentifikasi api pada citra digital berdasarkan ciri warna metode *K-Means Clustering* dan *Support Vector Machine (SVM)*. Dari rumusan masalah tersebut, penulis memutuskan untuk menggunakan metode simulasi

B. Conceptual Model

Pada tahap ini penulis membuat pemodelan yang akan digunakan untuk melakukan identifikasi api pada citra digital berdasarkan ciri warna menggunakan metode *K-Means Clustering* dan *Support Vector Machine (SVM)*.

C. Collection of Input Data

Langkah berikutnya, menentukan input yang akan diproses. Input yang akan diproses pada simulasi ini yaitu citra dengan format .jpg dan .png. Walaupun format gambar .tiff dan .bmp lebih unggul dalam menyimpan data gambar, namun penulis tidak menggunakan kedua format tersebut. Dikarenakan tidak tersedianya gambar yang bersifat opensource dalam kedua format tersebut.

Proses pengambilan data dilakukan dengan cara mencari sampel dataset gambar api dan gambar biasa. Selanjutnya dataset gambar tersebut di resize menjadi ukuran 150 * 150 piksel. Gambar api dan gambar biasa kemudian diproses untuk mendapatkan nilai ekstraksi warna RGB dan $YCbCr$ dengan menggunakan metode *K-Means Clustering*. Setelah nilai keduanya terkumpul dan dimasukkan dalam *database*, selanjutnya data citra diklasifikasikan menggunakan metode *Support Vector Machine*.

D. Modelling

Peneliti membuat model simulasi berupa model perhitungan manual metode ekstraksi warna $YCbCr$,

ekstraksi nilai warna *RGB* dan *YCbCr* menggunakan algoritma *K-Means Clustering*, dan algoritma *Support Vector Machine* yang diterapkan dalam penelitian ini.

E. Simulation

Pada tahap ini, peneliti menjalankan proses simulasi berdasarkan *conceptual model* atau model konseptual menggunakan teknik pengolahan citra yang telah dibuat pada tahap sebelumnya.

F. Verification and Validation

Tahap ini merupakan tahap di mana proses pemeriksaan data serta putusan akhir dari penulis untuk menilai data yang telah dihasilkan apakah layak atau tidak.

G. Experimentation

Tahap ini melakukan percobaan dengan menggunakan 80 data citra digital diluar data latih.

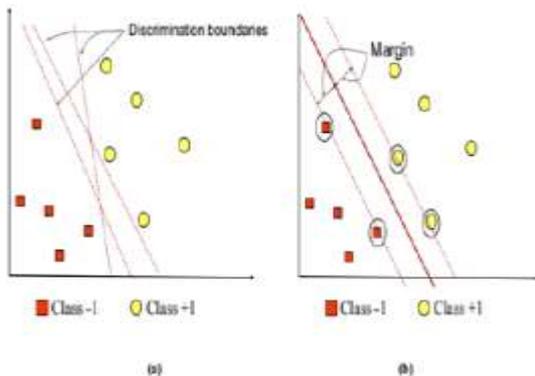
H. Output Analysis

Tahap output analysis merupakan tahap akhir dari metode simulasi. Pada tahapan ini, peneliti menganalisis output dari hasil percobaan yang telah dilakukan. Hasil analisis yang diuraikan berupa hasil akurasi yang diperoleh dalam penelitian sesuai dengan permasalahan utama.

I. Problem Formulation

Tahap ini melakukan pemodelan yang digunakan untuk melakukan pendeteksian api dalam citra digital berdasarkan ciri warna menggunakan metode *K-Means Clustering* dan *Support Vector Machine (SVM)*.

Gambar 1: Ilustrasi SVM Hyperplane



HASIL DAN PEMBAHASAN

Dalam mengidentifikasi api pada citra digital, ciri fitur didapatkan dari ekstraksi unsur warna *RGB* dan *YCbCr* dengan metode *K-Means Clustering* dan diklasifikasikan dengan menggunakan metode *Support Vector Machine (SVM)*. Pada proses ekstraksi unsur warna *RGB*, didapatkan tiga parameter, yaitu *Red*, *Green*, dan *Blue*. Begitu pula dengan unsur warna *YCbCr*, yang memiliki tiga parameter yaitu *Y*, *Cb*, dan *Cr*. Untuk normalisasi hasil ekstraksi untuk proses input, penulis

mengambil nilai mean dari parameter-parameter yang telah disebutkan.

Berikut tabel 1 adalah hasil nilai tengah variabel *Red (R)* dari perhitungan menggunakan *SVM (Support Vector Machine)*:

Tabel 1. Variabel Red (R) dari perhitungan SVM

R2	R1	Pred(R1)	Std. dev. on pred. (Mean)	Lower bound 95% (Mean)	Upper bound 95% (Mean)
2.328	36.429	25.712	3.364	17.954	33.470
2.402	29.094	25.680	3.353	17.949	33.412
9.050	33.105	22.812	2.954	15.999	29.625
18.545	18.367	18.714	4.537	8.252	29.175
5.580	22.926	24.309	2.988	17.418	31.200
3.305	27.724	25.291	3.221	17.863	32.719
2.649	29.031	25.574	3.315	17.930	33.218
0.598	6.748	26.459	3.668	18.000	34.918
0.960	21.108	26.303	3.600	18.001	34.604
30.595	9.834	13.514	7.859	-4.610	31.638

Dari tabel 1, nilai skala, *Prediction*, *Lower Bound*, dan *Upper Bound* merupakan nilai tengah dan nilai skala dari variabel.

Berikut hasil keseluruhan dari identifikasi api pada citra digital dengan menggunakan metode *K-Means Clustering* dan *Support Vector Machine (SVM)*:

A. Skenario 1

Dari hasil klasifikasi pada skenario yang menggunakan 20 citra digital hutan yang diambil saat siang hari terdapat 1 data yang salah identifikasi. Hasil dari uji akurasi untuk identifikasi api pada citra digital dengan menggunakan metode *K-Means Clustering* dan *Support Vector Machine* bisa dilihat dari tabel berikut:

Tabel 2. Matriks Konfusi Skenario 1

Citra hutan di siang hari	Terdeteksi Api	Terdeteksi Bukan Api	Jumlah
Api	10	0	10
Bukan Api	1	9	10
Jumlah	11	9	20
Akurasi	95 %		

Hasil perhitungan dari nilai akurasi tersebut dapat dilihat pada persamaan berikut ini:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$\frac{10 + 9}{10 + 9 + 1 + 0} \times 100\% = 95\%$$

Keterangan:

- *TP (True Positive)* : prediksi kelas positif & benar
- *TN (True Negative)* : prediksi kelas negatif & benar
- *FP (False Positive)* : prediksi kelas positif & salah
- *FN (False Negatif)* : prediksi kelas negatif & salah

B. Skenario 2

Dari hasil klasifikasi pada skenario yang menggunakan 20 citra hutan yang di ambil pada malam hari terdapat 3 data yang salah identifikasi. Hasil dari uji akurasi untuk identifikasi api pada citra digital dengan menggunakan metode *K-Means Clustering* dan *Support Vector Machine* bisa dilihat dari tabel berikut:

Tabel 3. Matriks Konfusi Skenario 2

Citra hutan di malam hari	Terdeteksi Api	Terdeteksi Bukan Api	Jumlah
Api	8	2	10
Bukan Api	1	9	10
Jumlah	9	11	20
Akurasi	85 %		

Hasil perhitungan dari nilai akurasi tersebut dapat dilihat pada persamaan berikut ini:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% = \frac{8+9}{8+9+1+2} \times 100\% = 85\%$$

Keterangan:

- *TP (True Positive)* : prediksi kelas positif & benar
- *TN (True Negative)* : prediksi kelas negatif & benar
- *FP (False Positive)* : prediksi kelas positif & salah
- *FN (False Negatif)* : prediksi kelas negatif & salah

C. Skenario 3

Dari hasil klasifikasi pada skenario yang menggunakan 20 citra bangunan yang di ambil pada siang hari terdapat 3 data yang salah identifikasi. Hasil dari uji akurasi untuk identifikasi api pada citra digital dengan menggunakan metode *K-Means Clustering* dan *Support Vector Machine* bisa dilihat dari tabel berikut:

Tabel 4. Matriks Konfusi Skenario 3

Citra bangunan di siang hari	Terdeteksi Api	Terdeteksi Bukan Api	Jumlah
Api	10	0	10

Bukan Api	1	9	10
Jumlah	11	9	20
Akurasi	95 %		

Hasil perhitungan dari nilai akurasi tersebut dapat dilihat pada persamaan berikut ini:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% = \frac{10+9}{10+9+1+0} \times 100\% = 95\%$$

Keterangan:

- *TP (True Positive)* : prediksi kelas positif & benar
- *TN (True Negative)* : prediksi kelas negatif & benar
- *FP (False Positive)* : prediksi kelas positif & salah
- *FN (False Negatif)* : prediksi kelas negatif & salah

D. Skenario 4

Dari hasil klasifikasi pada skenario yang menggunakan 20 citra bangunan yang di ambil pada malam hari terdapat 2 data yang salah identifikasi. Hasil dari uji akurasi untuk identifikasi api pada citra digital dengan menggunakan metode *K-Means Clustering* dan *Support Vector Machine* bisa dilihat dari tabel berikut:

Tabel 5. Matriks Konfusi Skenario 4

Citra bangunan di malam hari	Terdeteksi Api	Terdeteksi Bukan Api	Jumlah
Api	8	2	10
Bukan Api	0	10	10
Jumlah	18	2	20
Akurasi	90 %		

Hasil perhitungan dari nilai akurasi tersebut dapat dilihat pada persamaan berikut ini:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% = \frac{8+10}{8+10+0+2} \times 100\% = 95\%$$

Keterangan:

- *TP (True Positive)* : prediksi kelas positif & benar
- *TN (True Negative)* : prediksi kelas negatif & benar
- *FP (False Positive)* : prediksi kelas positif & salah
- *FN (False Negatif)* : prediksi kelas negatif & salah

KESIMPULAN DAN SARAN

Berdasarkan hasil pembahasan penelitian untuk identifikasi keberadaan api pada citra digital dengan ekstraksi unsur warna menggunakan *K-Means Clustering* dan klasifikasi menggunakan *Support Vector Machine* (SVM), maka kesimpulan yang dapat diambil adalah sebagai berikut:

1. Fitur RGB dan YCbCr dari citra digital hutan dan bangunan yang diambil pada siang dan malam hari akan diekstraksi dengan metode *K-Means Clustering*. Hasil ekstraksi tersebut akan digunakan sebagai data latih dan data uji, yang akan digunakan sebagai materi klasifikasi menggunakan *Support Vector Machine* (SVM).
2. Penggunaan ruang warna RGB dan YCbCr sebagai fitur yang digunakan untuk klasifikasi menggunakan *Support Vector Machine* (SVM) pada citra digital hutan dalam kondisi siang hari menghasilkan akurasi sebesar 95% dan 85% pada kondisi citra malam hari. Sedangkan pada citra digital bangunan pada siang hari dan malam hari menghasilkan akurasi sebesar 95%.
3. Berdasarkan percobaan, metode yang diusulkan tidak bekerja secara optimal pada citra Hutan di malam hari, yang menghasilkan 2 *False Negatif* dan 1 *False Positive*. Yang merupakan tingkat hasil *False* tertinggi dibandingkan dengan skenario lainnya.

Secara keseluruhan, penggabungan mode warna RGB dan YCbCr dalam mendeteksi api pada citra digital hutan dan bangunan memiliki tingkat keberhasilan >90%. Tingkat keberhasilan tersebut juga ditentukan dari metode klasifikasi yang penulis gunakan, yaitu *Support Vector Machine* (SVM). Hal itu pun juga bergantung dari data latih dan metode penggabungan data yang dipakai, yang dapat dikembangkan lagi oleh penelitian selanjutnya

Adapun saran peneliti untuk pengembangan penelitian selanjutnya sebagai berikut:

1. Dapat dikembangkan dengan menggunakan metode ekstraksi fitur lain, seperti fitur tekstur
2. Dapat dikembangkan menggunakan metode klasifikasi lainnya, serta dengan sampel citra yang lebih banyak
3. Dapat dikembangkan dengan mengklasifikasi api berdasarkan jenisnya, seperti tingkat besarnya api dan bahan bakar api
4. Dapat dikembangkan dengan mendeteksi api pada media video, sehingga api dapat dideteksi berdasarkan pola gerak api

DAFTAR PUSTAKA

- [1] UU No.24 2007
- [2] Mehaffey, James R., Joel L. Bert (1997) *Fire Protection. NIOSH Instructional Module. Ohio: U.S. Department of Health and Human Services*
- [3] Badan Standar Nasional. (2000)
- [4] Tawade, Hemangi. (2015). *Optimized Fire Detection Using Image Processing Based Techniques*
- [5] Rudz, S. (2013). *Investigation of Novel Image Segmentation Method Dedicated to Forest Fire*
- [6] Vezhnevets, Vladimir. (2003). *A Survey on Pixel-Based Skin Color Detection Technique*
- [7] Seebamrungsat, Jareerat. (2014). *Fire Detection in the buildings using Image Processing*
- [8] Jain, A. K. (2010). *Data Clustering: 50 years Beyond K-Means*
- [9] Lusiayanti, Desy. (2015). Perbandingan Metode Learning Vector Quantization (LVQ) dan Support Vector Machine (SVM) untuk Prediksi Penyakit Jantung Koroner
- [10] C. Emmy Premal (2014). *Image Processing Based Forest Fire Detection using YCbCr Colour Model*
- [11] Zuoning Wang (2017). *Research on Forest Flame Recognition Algorithm based on Image Feature*
- [12] K. N. Pushpalatha. (2012). *Iris Recognition System with Frequency Domain Features optimized with PCA and SVM Classifier*
- [13] Wei, Jin-Mao. (2010). *A novel measure for evaluating classifiers*
- [14] Solomon & Breckon. (2011). *Fundamentals of Digital Image Processing: A Practical Approach with Examples in Matlab*
- [15] Sargent, R.G. (2013). *Verification and Validation of Simulation Models. Journal of Simulation*
- [16] Gorunescu, Florin. (2011). *Data Mining: Concepts, Models and Techniques*