

**Deteksi dan Klasifikasi Konten Negatif pada Citra Berdasarkan
Kulit Manusia Dalam Ruang Warna HSV dengan Algoritma
*K-Nearest Neighbor***

Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika



Oleh :

Muhammad Rizki Akbar
NIM : 09021181722008

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2021**

LEMBAR PENGESAHAN SKRIPSI

Deteksi dan Klasifikasi Konten Negatif pada Citra Berdasarkan Kulit
Manusia Dalam Ruang Warna HSV dengan Algoritma *K-Nearest
Neighbor*

Oleh :

Muhammad Rizki Akbar

NIM : 09021181722008

Palembang, Juli 2021

Pembimbing I

Pembimbing II,

Nama Pembimbing I
NIP.

Nama Pembimbing II
NIP.

Mengetahui,
Ketua Jurusan Teknik Informatika

Rifkie Primartha, M.T.
NIP. 197706012009121004

DAFTAR ISI

Halaman

Deteksi dan Klasifikasi Konten Negatif pada Citra Berdasarkan Kulit Manusia Dalam Ruang Warna HSV dengan Algoritma <i>K-Nearest Neighbor</i> ..	i
LEMBAR PENGESAHAN SKRIPSI ..	i
DAFTAR ISI ..	ii
BAB I ..	3
PENDAHULUAN ..	3
1.1 Latar Belakang ..	3
1.2 Rumusan Masalah ..	4
1.3 Tujuan Penelitian ..	4
1.4 Manfaat Penelitian ..	4
1.5 Batasan Masalah ..	5
BAB II ..	6
KAJIAN LITERATUR ..	6
2.1 Pendahuluan ..	6
2.2 Landasan Teori ..	6
2.3 Penelitian Lain yang Relevan ..	8
2.4 Kesimpulan ..	8
BAB III ..	9
METODOLOGI PENELITIAN ..	9
3.1 Pendahuluan ..	9
3.2 Pengumpulan Data ..	9
3.3 Tahapan Penelitian ..	10
3.4 Metode Pengembangan Perangkat Lunak ..	12
3.5 Manajemen Proyek Penelitian ..	14
DAFTAR PUSTAKA ..	15

BAB I

PENDAHULUAN

1.1 Latar Belakang

Berdasarkan peraturan Menteri No 19 tahun 2014 pasal 4 ayat 1 jenis situs internet bermuatan negatif termasuk pornografi. Pornografi adalah penggambaran hal-hal cabul melalui tulisan, gambar, atau tontonan yang bertujuan mengeksploitasi seksualitas (Deny, 2010). Konten negatif berupa pornografi bisa ditemui dalam bentuk citra yang akan memberikan dampak buruk pada mereka yang belum siap menerimanya. **[1. Mengapa konten negatif ?] [1. Objek penelitian]**

Teknologi pendeteksian kulit sering digunakan untuk pengenalan manusia pada citra. Segmentasi warna kulit yang bisa digunakan salah satunya adalah menggunakan tipe citra HSV (Hue Saturation Value). Dari gambar dengan citra HSV tersebut kita bisa mendeteksi warna kulit manusia dengan menentukan range nilai warna kulit per pikselnya. **[2. Mengapa HSV?] Dengan begitu kita bisa mendapatkan warna kulit yang nantinya akan dideteksi (Oliveira, 2007). Jenis segmentasi ini dipilih karena citra yang mengandung konten negatif mudah dikenali karena didominasi warna kulit. [3. Mengapa deteksi konten negatif?] [2. Optimasi penelitian]**

Dalam pengkategorian citra apakah termasuk yang mengandung konten negatif atau bukan akan digunakan metode *K-Nearest Neighbor* **[4. Mengapa K-Nearest Neighbor?] [3. Metode yang dipilih]** dimana metode ini adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus

lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada. (Luthfi, Kusrini, 2009: 93).

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, masalah yang akan diselesaikan dalam penelitian ini adalah bagaimana mendeteksi dominasi warna kulit sebagai tolak ukur konten negatif pada citra dan mengkategorikannya menggunakan metode K-Nearest Neighbor.

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah membangun perangkat lunak yang dapat mendeteksi dominasi warna kulit sebagai tolak ukur konten negatif pada citra dan mengkategorikannya menggunakan metode K-Nearest Neighbor.

1.4 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah :

1. Perangkat lunak pendeteksi dan pengklasifikasian citra yang mengandung konten negatif dapat mempermudah menyeleksi gambar yang layak ditampilkan atau tidak.
2. Selanjutnya pengklasifikasian ini dapat diimplementasikan lebih lanjut ke dalam suatu sistem/mesin seperti peramban/ penjelajah internet.

1.5 Batasan Masalah

Batasan masalah dari penelitian ini adalah :

1. Harus terdapat objek manusia dalam citra yang dideteksi.
2. Akan selalu mendeteksi warna kulit walaupun dalam citra tidak ada objek manusia.

BAB II

KAJIAN LITERATUR

2.1 Pendahuluan

Pada bab ini membahas dasar-dasar teori yang akan digunakan dalam penelitian seperti definisi Citra, Ruang warna, *K-Nearest Neighbor* dan penelitian lain yang relevan.

2.2 Landasan Teori

2.2.1 Citra

Sebuah citra merupakan kumpulan piksel-piksel yang disusun dalam larik (array) dua dimensi. Sebuah piksel adalah sampel dari citra yang mengandung intensitas citra yang dinyatakan nilainya dalam bilangan bulat. Indeks baris dan kolom dari sebuah piksel (x,y) dinyatakan dalam bilangan bulat. Piksel dengan koordinat (0,0) terletak pada sudut kiri atas pada citra. Karakteristik operasi dalam pengolahan citra dapat dibedakan menjadi empat macam, yaitu operasi tingkat titik (point level), operasi tingkat lokal (local level), operasi tingkat global (global level), dan operasi tingkat obyek (object level) (Ahmad, 2005)

2.2.2 Ruang Warna

Ruang warna RGB menggunakan tiga komponen dasar yaitu merah (R), hijau (G), dan biru (B). Setiap piksel dibentuk oleh ketiga komponen tersebut. Model warna RGB biasa disajikan dalam bentuk kubus tiga dimensi, dengan warna merah, hijau, dan biru berada pada pojok sumbu (Kadir and Susanto, 2012). Ruang warna HSV dan HSL merupakan contoh ruang warna yang merepresentasikan warna seperti yang dilihat oleh mata manusia. H berasal dari kata “hue”, S berasal dari kata “saturation”, L berasal dari kata “Luminance”, I berasal dari kata “Intensity”, dan V berasal dari “value”. Ruang warna HLS terkadang disebut HLS, sedangkan HSV terkadang dinamakan HSB, dengan B berasal dari kata “Brightness”. Model HSV yang pertama kali diperkenalkan oleh A.R. Smith pada tahun 1978.

2.2.3 *K-Nearest Neighbor*

KNN adalah metode klasifikasi yang menentukan kategori berdasarkan mayoritas kategori pada tetangga terdekat. Jika D adalah sekumpulan data pelatihan maka ketika data uji d disajikan, algoritma akan menghitung jarak antara setiap data dalam D dengan data uji d . penghitungan jarak dilakukan dengan menggunakan euclidian distance. Kemudian k buah data dalam D yang memiliki jarak terdekat dengan d diambil. Himpunan k merupakan k -nearest neighbor. Selanjutnya kategori data uji d ditentukan berdasarkan label mayoritas kategori dalam himpunan k -tetangga terdekat (Farsiah dkk., 2013).

2.3 Penelitian Lain yang Relevan

Penelitian yang berkaitan dengan deteksi warna kulit pada citra dalam ruang warna HSV telah dilakukan oleh (Kolkur, 2017) dengan judul “*Human Skin Detection Using RGB, HSV and YCbCr Color Models*”. Dalam penelitian tersebut membandingkan 3 ruang warna yaitu RGB, HSV, dan Ycbcr. Hasilnya semua ruang warna bisa berkontribusi dalam mendeteksi warna kulit tetapi dalam prosesnya jauh lebih mudah bagi pengguna untuk mendapatkan warna yang diinginkan dengan memakai ruang warna HSV dibandingkan dengan menggunakan RGB.

Penelitian lain yang berkaitan dengan K-Nearest Neighbor sebagai algoritma klasifikasi gambar yang mengandung konten negatif telah dilakukan oleh (Nuraisha, 2017) dengan judul “*Implementation of K-NN Based on Histogram at Image Recognition for Pornography Detection*” Hasilnya gambar bisa diklasifikasikan menjadi gambar pornografi, semi-pornografi atau non-pornografi dengan akurasi 90%.

2.4 Kesimpulan

Bab II ini membahas hal-hal seperti Citra, Ruang warna, *K-Nearest Neighbor* dan penelitian lain yang relevan.

BAB III

METODOLOGI PENELITIAN

3.1 Pendahuluan

Bab ini menjelaskan tahapan penelitian yang akan dilaksanakan di dalam penelitian. Pada bab ini juga menjelaskan unit penelitian, pengumpulan data, dan metode pengembangan perangkat lunak yang digunakan. Tahapan penelitian dijadikan sebagai acuan pada setiap fase pengembangan dan memberikan sebuah solusi untuk rumusan masalah dan mencapai tujuan penelitian.

3.2 Pengumpulan Data

3.2.1 Jenis Data

Data yang digunakan dalam penelitian ini adalah human skin detection dataset yang didapat dari dataset CS Chan. Data disalin oleh peneliti dalam bentuk citra (.png) untuk kebutuhan penelitian. Data ini bersifat sekunder dan dipakai sebagai data latih dan data uji. Dataset ini terdiri dari citra yang diambil dengan jenis kamera, tingkatan warna, dan dalam kondisi cahaya yang berbeda-beda. Dataset terdiri dari 32 Facephoto (citra 1 orang) dan 46 Familyphoto (citra banyak orang).

3.2.2 Sumber Data

Dalam penelitian ini, peneliti menggunakan data yang didapat dari situs CS Chan dengan alamat “http://cs-chan.com/downloads_skin_dataset.html” yang menyediakan human skin detection dataset.

3.3 Tahapan Penelitian

Tahapan yang dilakukan dalam penelitian tugas akhir ini yaitu:

1. Pada tahap ini dilakukan studi terhadap konsep-konsep penunjang, metode yang digunakan, serta alat bantu dalam penelitian. Studi dilakukan dengan mengumpulkan informasi dari buku, jurnal, tugas akhir, dan lain-lain yang berkaitan dengan Citra dan metode klasifikasi, khususnya segmentasi citra dan algoritma klasifikasi, sehingga penulis mendapatkan dasar referensi yang kuat dalam menentukan metode yang tepat untuk menyelesaikan permasalahan yang diteliti.
2. Pada tahap ini dilakukan analisis dari algoritma dan skema yang dipilih penulis dalam penelitian. Selanjutnya penulis merumuskan masalah penelitian pada Bab I subbab 1.2, menentukan latar belakang masalah pada subbab 1.1, membuat tujuan dan manfaat penelitian pada subbab 1.3 dan 1.4, serta batasan masalah pada subbab 1.5.
3. Tahap pengembangan perangkat lunak (akan didokumentasikan pada bab IV).

4. Tahap pengujian.
5. Analisa pengujian (akan didokumentasikan pada bab V).
6. Pada tahap ini dilakukan penarikan kesimpulan dan saran berdasarkan hasil analisa pengujian. (akan didokumentasikan pada bab VI).

3.3.1 Kerangka Kerja

Segmentasi warna dengan deteksi warna HSV:

1. Tentukan citra RGB yang menjadi obyek deteksi, nilai warna HSV yang menjadi acuan (hasil proses pelatihan data) dan nilai toleransi HSV yang digunakan.
2. Transpose citra RGB ke HSV
3. Lakukan filter warna pada citra berdasarkan nilai acuan (T) dan nilai toleransi (tol).
4. Dengan x sebagai warna HSV pada piksel yang ada maka warna yang tidak termasuk dalam rentang $T - \text{tol} < x < T + \text{tol}$ diberi warna hitam.
5. Transpose kembali citra ke RGB, tampilkan hasil filter.

Kemudian pengidentifikasian citra yang mengandung konten negatif menggunakan algoritma *K-Nearest Neighbour (K-NN)*. Pada tahap ini melewati empat tahap yaitu menghitung jarak Euclidian, menghitung validasi data latih,

menghitung bobot data, dan identifikasi citra yang mengandung konten negatif berdasarkan dominasi warna kulit.

Setelah didapatkan nilai bobot dari setiap data latih maka akan dicocokkan dengan jumlah bobot dari data yang akan diuji. Dengan perbandingan nilai nya <20% bukan pornografi, 20-50% semi pornografi, > 50% pornografi. Tahap ini sebagai tahap klasifikasi citra apakah mengandung konten negatif atau tidak.

3.4 Metode Pengembangan Perangkat Lunak

Metodologi yang diterapkan dalam pengembangan perangkat lunak sebagai alat penelitian tugas akhir ini berorientasi pada objek menggunakan metode Rational Unified Process (RUP). Secara umum, langkah-langkah yang dilakukan pada pengembangan perangkat lunak ini adalah fase insepisi, fase elaborasi, fase konstruksi, dan fase transisi.

3.4.1 Fase Insepisi

Tahapan Pemodelan bisnis yang direncanakan oleh peneliti adalah penentuan user requirements dan fitur-fitur (fungsionalitas) yang dibutuhkan pada perangkat lunak. Pada tahapan pengumpulan kebutuhan, peneliti mengumpulkan data penelitian berupa dataset . Pada tahap analisis dan desain peneliti membuat diagram use case. Pada tahap implementasi, peneliti menyimpan dokumentasi user requirements, fungsionalitas perangkat lunak dan diagram use case. Pada tahap

pengujian, penulis akan memastikan user requirements dan fungsionalitas perangkat lunak telah valid.

3.4.2 Fase Elaborasi

Tahapan pemodelan bisnis yang direncanakan penulis adalah penentuan arsitektur perangkat lunak, desain basis data, dan desain antar muka sesuai dengan user equirements dan fungsionalitas perangkat lunak yang telah diperoleh. Peneliti dapat melengkapi user requirement, jika belum lengkap pada tahap pengumpulan kebutuhan.

3.4.3 Fase Konstruksi

Tahapan pemodelan bisnis yang direncanakan oleh peneliti adalah penentuan kelas-kelas yang dibutuhkan pada perangkat lunak. Pada tahap pengumpulan kebutuhan, ditentukan bahasa pemrograman yang digunakan untuk mengembangkan perangkat lunak yaitu Java. Kebutuhan lain dalam proses pengembangan perangkat lunak adalah perangkat keras dan perangkat lunak juga perlu diidentifikasi.

3.4.4 Fase Transisi

Tahapan pemodelan bisnis yang direncanakan oleh peneliti adalah peneliti membuat rencana atau skenario pengujian terhadap perangkat lunah. Peneliti menentukan tools pengujian yang diperlukan di tahap pengumpulan kebutuhan.

3.5 Manajemen Proyek Penelitian

Penjadwalan adalah perencanaan aktivitas penelitian dimulai dari tahap inisialisasi masalah sampai dengan tahap kesimpulan dari penelitian. Kegiatan yang berlangsung selama penelitian dapat dibuat dalam bentuk *Work Breakdown Structure (WBS)*.

DAFTAR PUSTAKA

- Ahmad, U., 2005, *Pengolahan Citra Digital & Teknik Pemrogramannya*, Graha Ilmu, Yogyakarta.
- Deny, W, 2010. *Hubungan antara pemanfaatan akses internet dengan sikap seks bebas pada remaja di Smk 1 Cokrominoto Surakarta*. Universitas Sebelas Maret.
- Farsiah, L, dkk, 2013, *Klasifikasi gambar berwarna menggunakan k-nearest neighbor dan support vector machine*, SNASTIKOM, Banda Aceh.
- Kadir, A. & Susanto, A. 2012, *Teori dan Aplikasi Pengolahan citra*, Penerbit Andi, Yogyakarta.
- Kusrini, Luthfi E. T, 2009, *Algoritma Data Mining*, Andi, Yogyakarta.
- Oliveira, V, 2007, *Skin Detection using HSV color space*, Computation Institute, Universidade Federal Fluminense – UFF – Niter.i, Brazil.
- Peraturan Menteri Komunikasi Dan Informatika Republik Indonesia Nomor 19, Tahun 2014 Tentang Penanganan Situs Internet Bermuatan Negatif. 2014. Kementerian Komunikasi dan Informatika, Jakarta.