

# FILTERISASI PORNOGRAFI PADA WEBSITE DENGAN METODE VIOLA JONES DAN DETEKSI WARNA KULIT

Nazori Agani<sup>1</sup>, Agustinus Sirumapea<sup>2</sup>

Program Studi Magister Ilmu Komputer, Program Pascasarjana, Universitas Budi Luhur

Jl. Raya Ciledug, Petukangan Utara, Kebayoran Lama, Jakarta Selatan 12260

Telp. (021) 5853753, Fax. (021) 5866369

<sup>1</sup>nazori.agani@gmail.com, <sup>2</sup>agust.rumapea@gmail.com

## ABSTRAK

*Pornografi dalam internet adalah salah satu hal yang dapat berdampak negatif bagi pemuda dan masyarakat . Untuk itu perlu untuk membuat aplikasi yang secara otomatis dapat memfilterisasi konten pornografi . Penelitian menggunakan penelitian terapan dalam bentuk prototipe untuk aplikasi deteksi gambar porno di halaman web . Metode yang digunakan adalah Viola dan Jones untuk deteksi objek dan dikombinasikan dengan deteksi kulit . Aplikasi ini diharapkan dapat mendeteksi gambar-gambar porno pada halaman web . Viola dan Jones metode yang digunakan untuk mendeteksi benda-benda sedangkan metode deteksi puting digunakan untuk mendeteksi warna kulit di sekitar puting warna adalah warna kulit manusia . Metode deteksi obyek dan deteksi gambar kulit manusia dan puting sana dengan tingkat efektifitas 90 % , tidak ada puting manusia efektifitas 94 % .*

*Kata kunci : Viola dan Jones algoritma , deteksi gambar-gambar porno , halaman web porno , deteksi warna kulit , perlindungan gambar-gambar porno .*

## I. PENDAHULUAN

Perkembangan teknologi komunikasi sangat pesat, itu dibuktikan dengan semakin banyaknya orang menggunakan internet untuk mencari informasi dan melakukan komunikasi. Hal ini membawa dampak negatif terhadap masyarakat terutama untuk usia anak-anak dan remaja karena mereka dengan mudah mengakses konten internet tanpa ada filterisasi.

Pornografi pada content internet adalah salah satu hal yang dapat berdampak negatif kepada masyarakat usia belia. Untuk melindungi mereka perlu ada aplikasi yang dapat memfilterisasi konten pornografi secara otomatis.

Deteksi gambar porno dari website merupakan masalah dalam *computer vision* dan *digital imaging process*. Menganalisa gambar porno sangat menarik perhatian untuk menelitinya. Sudah banyak penelitian dan pendekatan yang digunakan seperti Wavelet Image Pornography yang dikembangkan oleh James Ze Wang yang menggunakan deteksi kulit untuk menentukan telanjang apa tidak. Identifikasi gambar telanjang telah dirintis penelitiannya oleh Forsyth dengan pendekatan pengidentifikasian terhadap kulit dan pendefinisian gambar sebagai gambar telanjang.

Mendeteksi pixel warna kulit sepertinya adalah hal yang mudah tetapi dengan berbagai alasan cukup menantang untuk diteliti. Hal ini dapat dipengaruhi oleh penampilan kulit tergantung pada pencahayaan dan warna kulit yang berbeda-beda menurut ras. Pemisahan kulit dengan background image juga merupakan tantangan yang sulit. Ini bisa terjadi jika warna background menyerupai warna kulit.

Filterisasi pornography dengan deteksi obyek dan pendekatan deteksi kulit dengan mengkalkulasi proporsi kulit

pada satu gambar dengan lebih dulu mengekstrak image dari background. Mengkonversi image menjadi suatu color space tertentu yang sama dengan color space pada database warna kulit. Penentuan pixel warna kulit dengan membandingkan database sebagai perbandingan.

## II. LANDASAN TEORI

Dalam kamus Oxford Dictionaries pornografi memiliki definisi "printed or visual material containing the explicit description or display of sexual organs or activity, intended to stimulate sexual excitement"(bahan cetak atau visual yang berisi deskripsi eksplisit atau memamerkan organ atau aktivitas seksual, dimaksudkan untuk merangsang gairah seksual). Pornografi juga secara khusus merupakan konten yang berisi nudity (ketelanjangan) dan berbagai variasi adegan seks atau persetubuhan[1].

Definisi Pornografi menurut RUU Pornografi yang telah disahkan pada tanggal Pada 26 November 2008: "Pornografi adalah gambar, sketsa, ilustrasi, foto, tulisan, suara, bunyi, gambar bergerak, animasi, kartun, percakapan, gerak tubuh, atau bentuk pesan lainnya melalui berbagai bentuk media komunikasi dan/atau pertunjukan di muka umum, yang memuat kecabulan atau eksploitasi seksual yang melanggar norma kesusilaan dalam masyarakat.[2]. Secara khusus pornografi dapat dijelaskan sebagai konten yang berisi *nudity* (ketelanjangan) dan berbagai variasi adegan seks atau persetubuhan[3].

Sensual dalam keadaan atau memberi kesan telanjang bulat dilihat dari depan, samping atau belakang, penonjolan langsung alat-alat vital, payudara atau pinggul dan sekitarnya baik dengan penutup atau tanpa penutup; ciuman merangsang

antar pasangan sejenis atau berlainan jenis, baik antar muhrim ataupun non muhrim, atau antar manusia dengan binatang, antar binatang, atau antar manusia yang hidup dengan manusia yang telah meninggal dunia, gerakan atau bunyi dan atau desah yang memberi kesan persenggamaan atau percumbuan, gerakan masturbasi, lesbian, homoseksual, oral seks, sodomi, coitus interruptus, yang bertujuan untuk membangkitkan nafsu birahi dan atau yang menimbulkan rasa yang menjijikkan dan atau memuakkan dan atau yang memalukan bagi yang melihatnya dan atau mendengarnya dan atau menyentuhnya[4].

### A. Citra Digital

Citra digital adalah suatu citra  $f(x,y)$  yang memiliki koordinat spasial, dan tingkat kecerahan yang diskrit. Citra yang terlihat merupakan cahaya yang direfleksikan dari sebuah objek. Fungsi  $f(x,y)$  dapat dilihat sebagai fungsi dengan dua unsur. Unsur yang pertama merupakan kekuatan sumber cahaya yang melingkupi pandangan kita terhadap objek (*illumination*). Unsur yang kedua merupakan besarnya cahaya yang direfleksikan oleh objek ke dalam pandangan kita (*reflectance components*). Keduanya dituliskan sebagai fungsi  $i(x,y)$  dan  $r(x,y)$  yang digabungkan sebagai perkalian fungsi untuk membentuk fungsi  $f(x,y)$ . Fungsi  $f(x,y)$  dapat dituliskan dengan persamaan :

$$f(x,y) = i(x,y) * r(x,y) \quad (1)$$

di mana

$$0 < i(x,y) < \infty \text{ dan } 0 < r(x,y) < 1 \quad (2)$$

Citra digital merupakan suatu matriks yang terdiri dari baris dan kolom, dimana setiap pasangan indeks baris dan kolom menyatakan suatu titik pada citra. Nilai matriksnya menyatakan nilai kecerahan titik tersebut. Titik-titik tersebut dinamakan sebagai elemen citra, atau pixel (*picture element*) [5].

### B. Deteksi Gambar Porno

Secara garis besar algoritma deteksi gambar porno terbagi menjadi dua kategori[6]:

#### 1) Deteksi Berbasis Kulit

Algoritma deteksi berbasis kulit berfokus pada deteksi kulit. Yu-Chun Lin et.al.[7] melakukan deteksi gambar porno menggunakan algoritma *Support Vector Machine* (SVM) untuk pengenalan polanya dan distribusi warna kulit. Namun, algoritma ini memiliki kekurangan pada banyaknya gambar yang bukan porno tetapi mengandung banyak warna kulit dideteksi sebagai gambar porno. Begitu juga sebaliknya, gambar porno yang mengandung sedikit warna kulit dideteksi sebagai bukan gambar porno.

Johnson Agbinya et.al.[8] melakukan deteksi otomatis gambar porno online dengan melakukan segmentasi area kulit manusia dengan ruang warna HSV, jika area kulit pada sebuah gambar lebih dari 20% maka disimpulkan bahwa gambar tersebut

merupakan gambar porno. Algoritma ini pun memiliki kelemahan, yaitu pada gambar yang bukan porno dengan luasan warna seperti warna kulit akan dideteksi sebagai gambar porno.

Semua algoritma yang memanfaatkan deteksi kulit (*color histogram*, distribusi kromatik, informasi warna dan tekstur) mengalami permasalahan yang mendasar, yaitu tidak melakukan deteksi pada organ intim.

#### 2) Deteksi Berbasis Object

*Object detection* atau deteksi obyek merupakan suatu teknologi komputer yang berkaitan dengan *computer vision* dan *digital image processing* yang bertujuan mendeteksi suatu obyek (seperti manusia, bangunan, atau mobil) dalam gambar *digital* maupun *video digital*. Penelitian yang telah berhasil pada bidang *object detection* meliputi deteksi wajah dan deteksi pejalan kaki.

##### a) Haar Classifier

*Haar Classifier* merupakan pengklasifikasian fitur yang digunakan dalam algoritma *Viola and Jones*. Algoritma *Haar Classifier* membangun sebuah *boosted rejection cascade* yang akan membuat data *training* negatif, sehingga didapat suatu keputusan untuk menentukan data positif. *Haar Classifier* merupakan algoritma *supervised learning*, yaitu membutuhkan data *training* untuk dapat mendeteksi obyek-obyek tertentu. Untuk itu, *Haar Classifier* membutuhkan dua data set, yaitu data positif dan data negatif, di mana data positif merupakan data yang berisikan obyek yang akan dideteksi, sedangkan data negatif merupakan data yang berisikan obyek yang tidak akan dideteksi.

##### b) Pendeteksian Objek Viola Jones

Kerangka kerja pendeteksian objek Viola Jones adalah kerangka kerja pendeteksian objek pertama yang memberikan tingkat pendeteksian objek secara *realtime*. Kerangka kerja ini diusulkan pada tahun 2001 oleh Paul Viola dan Michael Jones dalam makalah mereka yang berjudul *Robust Real-Time Object Detection*. Makalah tersebut mendeskripsikan sebuah kerangka kerja pendeteksian objek visual, yang mampu mengolah citra dengan sangat cepat dan memberikan tingkat pendeteksian yang tinggi. Terdapat tiga kontribusi utama dalam kerangka kerja pendeteksian objek Viola Jones[9].

Kontribusi pertama adalah suatu representasi citra baru yang disebut dengan citra integral. Citra integral memungkinkan evaluasi fitur yang sangat cepat. Seperti Papageorgiou et al., Viola Jones menggunakan suatu kumpulan fitur yang mirip dengan fungsi *Haar Basis*. Agar dapat menghitung fitur-fitur tersebut dengan cepat pada berbagai skala, Viola Jones memperkenalkan suatu representasi citra integral. Citra integral dapat dihitung menggunakan sedikit operasi per piksel. Setelah citra integral dihitung, maka fitur Haar-like juga dapat dihitung pada berbagai skala dan tempat dalam waktu yang konstan.

Kontribusi ke dua dalam Viola Jones adalah metode untuk membangun suatu *classifier* dengan memilih sedikit fitur

penting menggunakan algoritma AdaBoost. Di dalam setiap citra *sub-window*, jumlah total fitur Haar-like sangat banyak, jauh lebih besar dari pada jumlah piksel. Agar klasifikasi dapat berjalan dengan cepat, proses pembelajaran harus menyisihkan sejumlah besar fitur yang ada, dan fokus pada sejumlah kecil fitur-fitur penting. Penyeleksian fitur pada metode Viola Jones didapatkan dari modifikasi sederhana terhadap AdaBoost. Setiap tahapan pada proses *boosting*, yang memilih sebuah *classifier* lemah baru, dapat dilihat sebagai proses penyeleksian fitur.

AdaBoost menyediakan sebuah algoritma pembelajaran yang efektif dan batas-batas yang kuat terhadap kinerja generalisasi. Kontribusi ke tiga adalah metode untuk menggabungkan *classifier* yang lebih kompleks secara berturut-turut di dalam suatu struktur *cascade* (bertingkat) yang meningkatkan kecepatan detektor secara dramatis dengan memfokuskan perhatian pada daerah-daerah yang menjanjikan di dalam citra. Pemrosesan yang lebih kompleks hanya akan dijalankan pada daerah-daerah yang menjanjikan ini.

*Sub-windows* yang tidak ditolak oleh *classifier* terdahulu akan diproses oleh sederetan *classifier* lainnya, di mana setiap tingkatan akan sedikit lebih kompleks dibandingkan tingkatan sebelumnya. Jika suatu *classifier* menolak *sub-window*, maka tidak akan ada proses lebih lanjut untuk *sub-window* tersebut.

*Cascade* deteksi wajah memiliki 32 *classifiers*, dengan total lebih dari 80.000 operasi. Namun, struktur bertingkat ini mampu mendeteksi dalam waktu yang sangat cepat. Pada set data yang sulit, berisi 507 wajah dan 75 juta *sub-window*, wajah terdeteksi menggunakan sekitar 270 instruksi mikroprosesor per *sub-window*. Sebagai perbandingan, sistem ini lebih cepat 15 kali dari pada implementasi sistem pendeteksian yang dibuat oleh Rowley et al. Walaupun dapat dilatih untuk mendeteksi variasi dari kelas-kelas objek, pada dasarnya, sistem pendeteksian objek Viola Jones dimotivasi oleh permasalahan pendeteksian wajah.

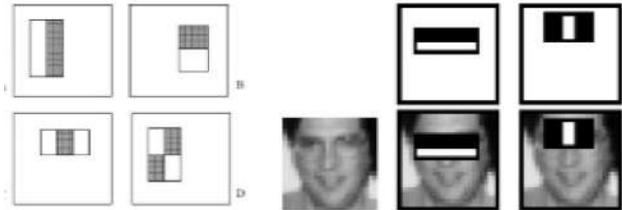
**c) Fitur Haar-Like**

Fitur Haar-like adalah fitur citra digital yang digunakan di dalam pengenalan objek. Sejumlah fitur Haar-like mewakili wilayah persegi pada citra dan menjumlahkan semua piksel pada daerah tersebut. Jumlah yang didapatkan digunakan untuk mengkategorisasikan citra. Pada citra *sample* 20x20, terdapat lebih dari 40.000 fitur, yang harus diseleksi menjadi 200 fitur baik.

Sistem pendeteksian objek Viola Jones mengklasifikasikan citra berdasarkan nilai dari fitur-fitur sederhana. Ada beberapa alasan untuk menggunakan fitur dan bukan piksel. Salah satu alasan utamanya adalah bahwa fitur dapat meng-*encode* pengetahuan domain *ad-hoc* yang sulit dipelajari menggunakan data latih dengan jumlah yang terbatas. Alasan lainnya adalah bahwa sistem berbasis fitur beroperasi lebih cepat dari pada sistem berbasis pixel.

Fitur sederhana yang digunakan mirip dengan fitur dasar Haar yang digunakan oleh Papageorgiou et al. Viola Jones menggunakan tiga jenis fitur. Nilai dari sebuah fitur dua-

persegi adalah selisih jumlah piksel di antara kedua daerah persegi tersebut. Setiap daerah tersebut memiliki bentuk serta ukuran yang sama, dan berbatasan secara horizontal ataupun secara vertikal. Nilai dari sebuah fitur tiga-persegi adalah hasil pengurangan antara jumlah piksel pada kedua persegi yang berada di luar dengan jumlah piksel pada persegi yang berada di tengah. Sedangkan nilai dari sebuah fitur empatpersegi adalah selisih antara setiap pasangan diagonal persegi [9].



Gambar 1. Fitur persegi Haar-like

**d) Image Integral**

Fitur persegi dapat dihitung dengan cepat menggunakan suatu representasi citra yang disebut citra integral. Nilai citra integral pada titik(x,y) adalah jumlah piksel yang ada di atas dan di sebelah kiri titik tersebut.

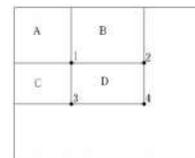
$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

Di mana *ii(x,y)* adalah citra integral dan *i(x,y)* adalah citra asli.

$$S(x,y) = s(x,y-1) + i(x,y) \tag{4}$$

$$ii(x,y) = s(x,y-1) + i(x,y) \tag{5}$$

Di mana *S(x,y)* adalah jumlah baris kumulatif, *s(x-1,y)=0* dan *ii(-1,y)=0*.



Gambar 2. Nilai citra integral pada titik (x,y) adalah jumlah semua piksel yang berada di atas dan di kiri titik tersebut

Dengan menggunakan citra integral, jumlah dari setiap persegi dapat dihitung dengan menggunakan empat referensi array. Perbedaan antara jumlah dua persegi dapat dihitung dengan delapan referensi. Fitur dua persegi dapat dihitung dengan enam referensi array, delapan untuk fitur tiga persegi, dan sembilan untuk fitur empat persegi [10].

1	1	1
1	1	1
1	1	1

1	2	3
2	4	6
3	6	9

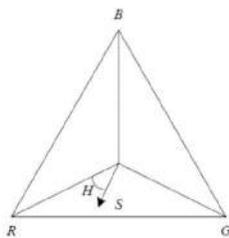
Gambar 3. Citra integral

**C. Transformasi Warna RGB ke IHS**

Meskipun basis *RGB* bagus untuk menampilkan informasi warna, tetapi ia tidak cocok untuk beberapa aplikasi pemrosesan citra. Pada aplikasi pengenalan objek, lebih mudah mengidentifikasi objek dengan perbedaan *hue*-nya dengan cara memberikan nilai ambang pada rentang nilai-nilai *hue* (panjang gelombang spektrum) yang melingkupi objek. Masalahnya, bagaimana melakukan pengambangan pada ruang warna *RGB* dan apa rumus untuk mengaplikasikannya? Masalah ini lebih mudah dipecahkan bila nilai *RGB* dikonversi ke nilai *intensity* (*I*), *hue* (*H*), dan *saturation* (*S*).

Aplikasi yang lain misalnya dalam pemampatan citra. Melakukan pemampatan secara terpisah pada setiap nilai *R*, *G*, dan *B* tidak disarankan, karena data yang dimampatkan 3 kali lebih banyak dan waktu pemampatannya 3 kali lebih lama daripada waktu pemampatan citra skala-abunya. Pemampatan citra berwarna lebih relevan bila warna *RGB*-nya dikonversikan ke *IHS* karena algoritma pemampatan pada citra skala-abu dilakukan pada komponen *I*, sedangkan nilai *H* dan *S* dikodekan dengan cara yang lain dengan sedikit atau sama sekali tidak ada degradasi.

Model warna *IHS* merepresentasikan warna dalam terminologi *intensity*, *hue*, dan *saturation*. Dari diagram kromatisitas, buatlah segitiga yang menghubungkan tiga warna pokok *red*, *green*, *blue*. Titik-titik pada segitiga menyatakan warna yang dihasilkan dari pencampuran warna titik sudut, sedangkan titik-titik di dalam segitiga menyatakan warna yang dapat dihasilkan dengan mengkombinasikan tiga warna titik sudut. Titik tengah segitiga menyatakan warna putih, yaitu pencampuran warna pokok dengan fraksi yang sama



Gambar 4. Segitiga HIS

Komponen *RGB* dari citra berwarna dapat dikonversikan ke model warna *IHS*. Dengan mengasumsikan komponen *RGB* telah dinormalisasikan ke 1, maka *I* dihitung dengan rumus:

$$I = \frac{1}{3}(R + G + B) \tag{6}$$

Persamaan (6) di atas sering digunakan untuk mengubah citra berwarna menjadi citra skala abu-abu.

$$S = 1 - \frac{3}{R + G + B} \min(R, G, B) \tag{7}$$

Nilai *H* dihitung dengan rumus: Nilai *S* dihitung dengan rumus:

$$H = \cos^{-1} \frac{2R - G - B}{2\sqrt{(R - G)^2 + (R - B)(G - B)}} \tag{8}$$

Alternatif lain mengubah model *RGB* ke model *IHS* adalah sebagai berikut. Konversi dari model *RGB* ke model *IHS* dilakukan dalam dua tahap. Tahap pertama adalah merotasikan koordinat *RGB* ke sistem koordinat (*I*, *V1*, *V2*) dengan transformasi:

$$H = \tan^{-1}(V_2/V_1) \tag{9}$$

$$S = (V_1^2 + V_2^2)^{1/2} \tag{10}$$

Nilai *H* adalah dalam selang [0,2p] atau setara dengan [0, 360°]. Transformasi dari model *IHS* ke model *RGB* dapat dilakukan dengan prosedur balikan:

$$V_1 = S \cos(H) \tag{11}$$

$$V_2 = S \sin(H) \tag{12}$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} \sqrt{3}/3 & 0 & 2/\sqrt{6} \\ \sqrt{3}/3 & 1/\sqrt{2} & -1/\sqrt{6} \\ \sqrt{3}/3 & -1/\sqrt{2} & -1/\sqrt{6} \end{bmatrix} \begin{bmatrix} I \\ V_1 \\ V_2 \end{bmatrix} \tag{13}$$

Dengan transformasi *RGB* ke *IHS*, maka algoritma pemrosesan citra yang semula untuk citra skala-abu dapat diterapkan pada komponen *intensity*, sedangkan algoritma segmentasi citra dapat dilakukan pada komponen *H*. Transformasi citra dari basis *RGB* ke basis *IHS* dilakukan sebelum pemrosesan citra. Citra yang sudah diproses dapat dikonversikan kembali ke basis *RGB* untuk tujuan *display*.

National Television Systems Committee (*NTSC*) menggunakan model *YIQ* untuk mentransformasikan model *RGB* ke *IHS*, yang dalam hal ini *Y* menyatakan *intensity*, *I* menyatakan *hue*, dan *Q* menyatakan *saturation*.

#### D. Deteksi gambar porno dengan Algoritma Rigan Ap-apid

Algoritma deteksi gambar porno ini dikembangkan oleh Rigan Ap-apid dari De La Salle University, Manila, Filipina mendeteksi gambar porno dengan tingkat keberhasilan yang cukup tinggi yaitu 90%.

Algoritma yang diajukan oleh Rigan Ap-apid adalah :

Langkah awal untuk mendeteksi citra pornografi dengan Nudity Detection Algorithm adalah sebagai berikut:

1. Menghitung prosentase pixel kulit dari keseluruhan pixel citra.
2. Membentuk skin region dari piksel-piksel kulit yang telah diklasifikasi. Skin region merupakan kumpulan piksel-piksel kulit yang saling berhubungan satu sama lain.
3. Menghitung jumlah skin region yang didapat.
4. Mengidentifikasi tiga buah skin region yang paling luas.
5. Menghitung prosentase skin region yang paling besar terhadap ukuran citra.
6. Mencari piksel kulit yang terletak paling kiri, paling kanan, paling atas, dan paling bawah dari ketiga skin region yang paling besar. Lalu hubungkan titik-titik tersebut sehingga membentuk sebuah segiempat saling berhubungan yang disebut dengan bounding polygon.
7. Hitung luas bounding polygon.
8. Hitung jumlah piksel kulit yang terdapat pada bounding polygon.
9. Hitung prosentase piksel kulit pada bounding polygon terhadap luas bounding polygon tersebut.
10. Hitung intensitas rata-rata dari setiap piksel pada bounding polygon.

Kemudian citra yang ingin dideteksi diklasifikasikan menurut aturan berikut :

1. Jika prosentase jumlah piksel kulit terhadap jumlah keseluruhan piksel citra kurang dari 15%, maka citra tersebut bukan citra pornografi.
2. Jika jumlah piksel kulit pada skin region yang paling besar kurang dari 35% dari jumlah semua piksel kulit, dan jumlah piksel kulit pada skin region kedua terbesar kurang dari 30% dari jumlah semua piksel kulit, dan jumlah piksel kulit pada skin region ketiga terbesar kurang dari 30% dari jumlah semua piksel kulit, maka citra tersebut bukan citra pornografi.
3. Jika jumlah piksel kulit pada skin region yang paling besar kurang dari 45% dari jumlah semua piksel kulit, maka bukan citra pornografi.
4. Jika jumlah piksel kulit kurang dari 30% dari jumlah keseluruhan piksel citra dan jumlah piksel kulit pada bounding polygon kurang dari 55% dari luas polygon, maka citra tersebut bukan citra pornografi.
5. Jika jumlah skin-region lebih dari 60 dan intensitas rata-rata pada polygon kurang dari 0,25, maka citra tersebut bukan citra pornografi.
6. Jika tidak memenuhi salah satu kriteria yang telah disebutkan, maka citra tersebut adalah citra pornografi.

#### E. Tinjauan Studi

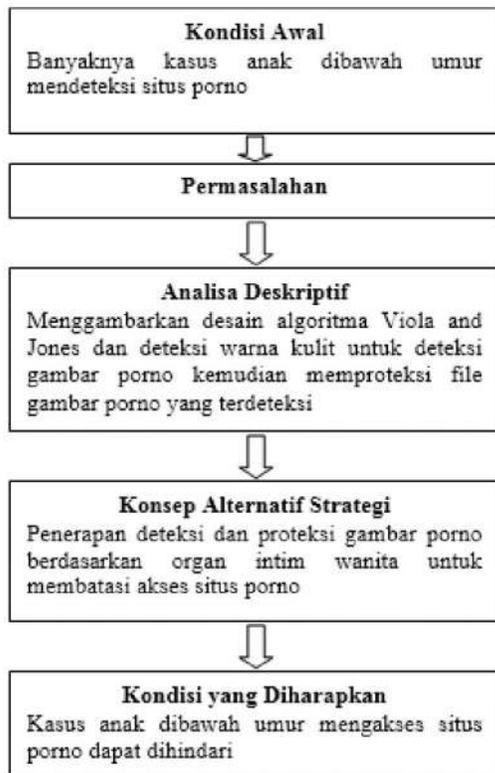
Tinjauan studi yang dijadikan acuan dalam melakukan penelitian ini mengacu pada penelitian-penelitian yang menggunakan algoritma deteksi obyek *Viola and Jones* pada gambar digital. Berikut adalah ringkasan dari beberapa penelitian terkait:

1. Muhammad Fikri Hidayattullah dan Yustia Hapsari, 2013. *Automatic Nipple Detection* Pada Citra Pornografi Menggunakan Algoritma *Viola and Jones* Berbasis *AdaBoost* Untuk *Feature Selection* *Viola and Jones* 897 gambar positif (mengandung obyek puting), 1836 gambar negatif (tidak mengandung obyek puting). Gambar positif terdiri dari 731 gambar latih dan 166 gambar uji. Gambar negatif terdiri dari 1756 gambar latih dan 80 gambar uji. Gambar untuk pengujian tidak sama dengan untuk pelatihan. Puting payudara yang digunakan sebagai *data training* maupun *data testing* memiliki bentuk, ukuran, warna, tekstur, posisi dan pencahayaan yang berbeda-beda dikarenakan bentuk payudara wanita sangat . Menambahkan deteksi organ intim (puting payudara) karena deteksi gambar porno berdasarkan warna kulit saja dianggap kurang efektif. Deteksi hanya dilakukan pada gambar payudara yang bersifat *frontal*. Sehingga, semua gambar payudara yang tidak menghadap lurus ke depan (*frontal*) tidak menjadi obyek deteksi dan tidak dihitung sebagai *false negative*. Algoritma *Viola and Jones* untuk deteksi gambar porno dengan memanfaatkan salah satu organ intim, yaitu puting payudara terbukti akurat untuk mengenali obyek *nudity*. Akan tetapi, masih banyak area bukan puting yang dianggap sebagai puting karena memiliki kesamaan warna dan bentuk seperti mata dan pusar. Ke depannya algoritma ini perlu digabungkan dengan algoritma lain, misalnya saja deteksi kulit agar menghasilkan deteksi yang lebih akurat. Rata-rata akurasi yang dihasilkan dari penelitian mereka sebesar 83,75%, *precision rate* sebesar 88,20%, dan *recall rate* sebesar 90,21%.
2. Jerry Pratama Hakim, 2012. Perancangan Program Aplikasi Pendeteksian Bagian Tubuh Manusia Dengan Metode *Viola and Jones* 100 sampel gambar positif dari kamera BB 9760 5MP, dan 200 sampel gambar negatif dari google.com. Deteksi bagian tubuh manusia (wajah, mata, hidung, telinga kiri dan kanan, tubuh bagian atas, dan tangan kiri) dari sembarang gambar *digital* untuk program filter pornografi. Deteksi bagian tubuh manusia (wajah, mata, hidung, telinga kiri dan kanan, tubuh bagian atas, dan tangan kiri) dari sembarang gambar *digital* untuk program filter pornografi.

3. Huicheng Zheng, 2004. Blocking Adult Images Based on Statistical Skin Detection Database gambar terdiri dari 1. 297 Gambar porno. Deteksi warna kulit dan menghitung prosentase warna kulit dan menggunakan deteksi untuk wajah manusia. Tingkat keberhasilan 83, 2%.

### III. KERANGKA PEMIKIRAN

Pada penelitian ini, kerangka pemikiran yang akan digunakan untuk menyelesaikan rumusan masalah dapat dilihat pada gambar di bawah ini:



Gambar 5. Kerangka Pemikiran Deteksi dan Proteksi Gambar Porno

Berikut ini adalah penjelasan dari kerangka pemikiran deteksi dan proteksi gambar porno dari gambar di atas:

#### a. Kondisi awal

Banyaknya kasus pelecehan seksual yang terjadi terutama kepada anak-anak harus dihindari, salah satunya dengan membatasi penyebaran gambar porno. Deteksi gambar porno berdasarkan warna kulit saja dianggap kurang efektif untuk membatasi penyebaran gambar porno.

#### b. Permasalahan

Rumusan identifikasi masalah pada penelitian ini yaitu obyek deteksi gambar porno berdasarkan organ intim wanita meliputi puting payudara dan menghitung statistik warna kulit disekitar obyek. Selain itu dari gambar yang berhasil diproteksi kemudian dilakukan proteksi sehingga diharapkan penyebaran gambar porno dapat dibatasi.

#### c. Analisis Deskriptif

Pada analisis deskriptif Penulis mengusulkan desain algoritma *Viola and Jones* yang akan digunakan untuk mendeteksi adanya obyek-obyek dari suatu gambar berupa organ intim wanita yang meliputi puting payudara dari berbagai sudut gambar. Untuk memastikan deteksi gambar benar ditambahkan deteksi warna kulit.

#### d. Konsep Alternatif Strategis

Pada konsep alternatif strategi desain algoritma *Viola and Jones* yang diusulkan akan digambarkan menggunakan *flowchart* kemudian diimplementasikan ke dalam bentuk model prototipe aplikasi yang akan dibangun menggunakan MATLAB untuk deteksi gambar porno dan php, javascript untuk memberikan peringatan atau keluar dari halaman web yang dikunjungi.

#### e. Kondisi Akhir yang Diharapkan

Gambar porno yang mengandung puting diharapkan akan diberi peringatan atau keluar dari halaman web yang dikunjungi .

### Hipotesis

- Diduga aplikasi mendeteksi gambar yang mengandung gambar porno dan aplikasi akan mendeteksinya.
- Diduga jika ada kecocokan dengan fitur dari data training maka gambar akan diblok..
- Diduga jika proporsi warna kulit gambar dominan maka dideteksi gambar sebagai gambar porno dan diblok

## IV. METODOLOGI RANCANGAN PENELITIAN

### A. Metode Penelitian

Menggunakan penelitian terapan dalam bentuk prototipe untuk aplikasi deteksi gambar porno pada halaman web. Metode yang digunakan adalah *viola and jones* untuk deteksi objek dan digabung dengan deteksi kulit. Aplikasi ini diharapkan dapat mendeteksi gambar porno pada halaman web.

### B. Metode Pengumpulan Data

Metode mengumpulkan data dilakukan dengan pengamatan pada objek penelitian dan studi dokumentasi yang

berhubungan dengan aplikasi deteksi gambar porno pada II. website. Gambar porno dikumpulkan dari berbagai situs internet seperti gambar puting, gambar yang mengandung pornografi dan gambar yang tidak mengandung pornografi.

**C. Sumber Data**

Data menggunakan dari berbagai pustaka yaitu buku, paper, tesis dan sumber-sumber dari internet.

**D. Teknik Analisis Data**

Teknik analisis data dalam penelitian tesis ini menggunakan pendekatan kuantitatif. Analisis data kuantitatif dilakukan dengan cara menganalisis proses training data set , penentuan positif image dan negatif image. Kemudian menganalisa proses penentuan gambar sebagai gambar porno atau tidak. Ujicoba dilakukan pada beberapa gambar dengan berbagai situasi yaitu gambar yang mengandung puting, tidak mengandung puting tetapi masih tergolong gambar porno dan gambar bukan porno.

**E. Formulasi Hipotesis**

Hipotesis merupakan prediksi dan kesimpulan sementara mengenai tingkat akurasi dalam pendeteksian gambar pada halaman web yang mengandung puting dengan metode viola jones dan deteksi warna kulit.

**F. Desain Algoritma Deteksi Pornografi**

Mempelajari teknik pemrograman Matlab, serta mengumpulkan dan mempelajari algoritma *Viola and Jones* dan metode deteksi warna kulit yang akan digunakan dalam perancangan model prototipe aplikasi deteksi dan proteksi gambar porno.

Mengelompokkan cara kerja model prototipe aplikasi ke dalam 2 bagian utama, modul apa saja yang diperlukan, merancang model prototipe aplikasi deteksi dan proteksi halaman web yang mengadung gambar porno. Perancangan meliputi: perancangan pembelajaran gambar (*training image*) dengan memanfaatkan fitur (puting payudara) perancangan deteksi gambar porno berdasarkan fitur-fitur yang ditentukan, dan perancangan proteksi gambar porno.

Model aplikasi yang dirancang secara garis besar adalah:

**I. Deteksi Gambar Porno**

Pengguna mengunjungi salah satu halaman web. Kemudian aplikasi akan mendeteksi apakah ada gambar porno yang mengandung feature puting. Proses deteksi gambar porno berdasarkan puting dan deteksi warna kulit . Jika gambar terdeteksi gambar puting maka dideteksi itu gambar porno dan jika terdeteksi puting akan dilanjutkan dengan deteksi warna kulit. Jika proporsi warna kulit tinggi maka akan dideteksi gambar itu porno.

**Proteksi Gambar Porno**

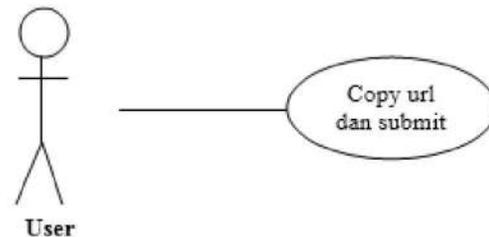
Halaman web yang terdeteksi gambar porno diberi peringatan , diblok halaman webnya dan keluar dari halaman web. Algoritma Deteksi pornografi dengan metode deteksi warna kulit:

1. Mendeteksi piksel kulit berwarna dalam gambar.
2. Cari daerah kulit berdasarkan deteksi pixel kulit.
3. Menganalisis daerah kulit untuk menentukan ketelanjangan atau porno
4. Hitung persentase piksel kulit relatif terhadap ukuran gambar.
5. Jika persentase ukuran piksel warna kulit lebih dari 15% dari original image gambar porno.

**G. Desain Dan Analisis Berorientasi Objek**

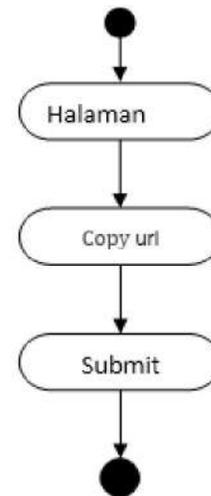
**1. Use Case Diagram**

Use case diagram menggambarkan interaksi user dengan aplikasi.



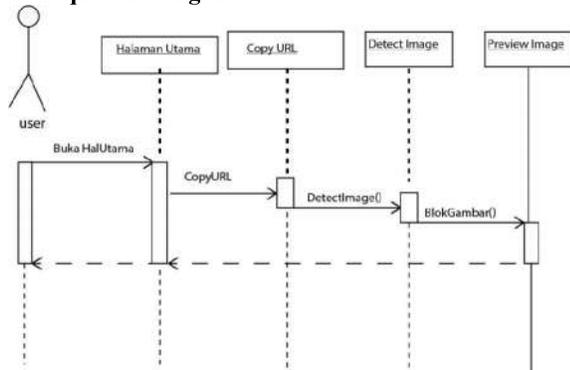
Gambar 6. Use Case Diagram Deteksi Pornografi

**2. Activity Diagram**



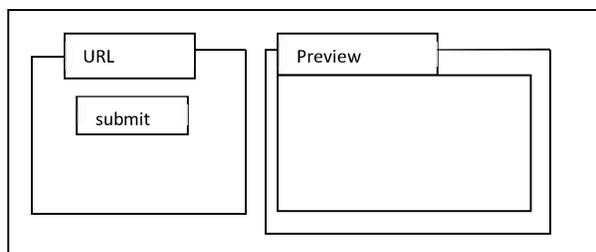
Gambar 7. Activity Diagram Deteksi Pornografi

### 3. Sequence Diagram



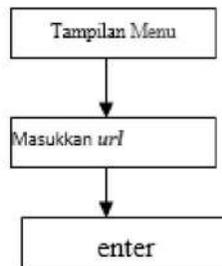
Gambar 8. Activity Diagram Deteksi Pornografi

### 4. Rancangan Layar



Gambar 9. Rancangan Menu Deteksi Pornografi

### 5. Alur Sistem yang Dikembangkan



Gambar 10. Alur deteksi pornografi

### H. Pengujian dan Analisis

Pengujian terhadap aplikasi yang telah dirancang dilakukan dengan menjalankannya pada sistem operasi Windows. Pengujian dilakukan dengan browser mozilla firefox dan mengunjungi beberapa halaman web yang mengandung pornografi. Kemudian dari halaman web yang dikunjungi dianalisa prosentase keakuratan aplikasi dalam mendeteksi gambar porno.

### I. Penarikan Kesimpulan

Kesimpulan dibuat berdasarkan hasil pengujian terhadap aplikasi yang dirancang. Aplikasi dinyatakan berhasil apabila mampu mendeteksi adanya gambar porno pada halaman web dan memberikan peringatan .

### V. PEMBAHASAN HASIL PENELITIAN

Model deteksi gambar porno berdasarkan organ intim wanita yaitu puting menggunakan algoritma Viola-Jones dan metode deteksi kulit. Adapun tujuan penelitian ini diharapkan dapat membantu orangtua dalam mengawasi mana gambar yang layak dilihat maupun mana gambar yang tidak layak dilihat oleh anaknya (*parental control*). Model diimplementasikan menggunakan Matlab.

#### A. Persiapan Data

Sampel yang digunakan pada penelitian ini, sebagaimana yang tertulis pada bab sebelumnya, menggunakan 2 (dua) set atau kelompok data sampel. Pertama, kelompok data positif, berisi gambar-gambar yang di dalamnya memuat obyek organ intim wanita, yaitu: puting payudara dengan berbagai bentuk maupun ukuran. Kedua, kelompok data negatif, berisi gambar-gambar yang tidak memuat obyek organ intim wanita. Masing-masing kelompok dapat terdiri dari gambar berwarna maupun gambar grayscale.

##### 1. Kelompok Data Positif

Kelompok data positif berisi obyek yang akan dideteksi. Pengambilan sample untuk data positif disesuaikan dengan bentuk dan sudut obyeknya. Semakin banyak sample akan mempengaruhi keakuratan deteksi gambar.

##### 2. Kelompok Data Negatif

Kelompok data negatif berisi obyek yang tidak akan dideteksi. Pengambilan sample untuk data negatif disesuaikan dengan bentuk dan sudut obyeknya. Semakin banyak sample akan mempengaruhi keakuratan deteksi gambar.

#### B. Langkah-langkah di Matlab

1. Persiapan Data Training Set  
Dalam proses ini kita menggunakan aplikasi *Adobe photoshop* sampel data set positif dan set data negatif.
2. Membuat sampel Data negative dan sampel data positif  
Sampel Data negative adalah sekumpulan gambar yang tidak mengandung obyek yang akan dideteksi. Data positif adalah sekumpulan gambar yang mengandung obyek yang akan dideteksi.



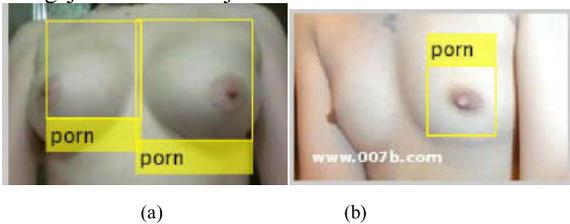
Gambar 11. Gambar Contoh data positif



Gambar 12. Gambar Contoh data negatif

Pada step *training* ini akan dihasilkan sebuah *cascade classifier* yang berisi sejumlah stage yang didalamnya terdapat kumpulan data classifier. Cascade classifier dihasilkan dengan menggunakan aplikasi *Train Cascade Detector* dari *Cascade Train GUI*. Berkas classifier ini berupa XML.

### 3. Pengujian Deteksi Objek Metode Viola Jones



Gambar 13. Hasil experimental dengan metode viola jones untuk deteksi puting payudara

Kelemahan pada pengujian metode ini adalah untuk payudara yang tidak frontal tidak dapat terdeteksi dengan baik. Hal itu terbukti pada gambar 13 puting yang sebelah kiri tidak terdeteksi.

### 4. Deteksi Kulit

Dalam sub bagian berikut, model multi-warna statistik untuk deteksi kulit yang menggunakan RGB, Normalized RGB, Y CrCb dan HSL.



Gambar 14. Hasil segmentasi skin color 1

Deteksi kulit untuk meningkatkan keakuratan deteksi gambar porno dan mengatasi ketidakakuratan pada metode objek viola jones. Dari pengujian kelemahan metode ini adalah jika *background* dari objek warnanya menyerupai kulit

sehingga segmentasi kulit dan tidak kulit tidak akurat. Hal ini dibuktikan pada gambar 15.



Gambar 15. Hasil segmentasi skin color 2

### 5. Pengujian Kecepatan Deteksi Gambar

Deteksi dengan metode viola jones lebih cepat dibanding dengan deteksi kulit yaitu 0.256 s untuk Viola Jones dan 1.979 s untuk deteksi kulit.

### C. Hasil Pengujian

Performansi Breast Detection dan Skin Detection menurut hasil eksperimen yang dilakukan pada beberapa gambar sebagai berikut :

Tabel 1. Hasil *Experiment*

Jenis Gambar	Terdeteksi puting	Tidak terdeteksi puting	Efektivitas
Manusia + puting (ada)	120	12	90%
Manusia - puting	5	80	94
Bukan Manusia	2	95	98

Dilihat dari tabel diatas terdapat beberapa kegagalan dalam mendeteksi gambar porno diakibatkan gambar puting kurang jelas karena posisinya yang kurang *frontal*. Pada gambar yang tidak mengandung puting tetapi terdeteksi puting terjadi pada gambar yang mempunyai kemiripan dengan puting.

Pada jenis gambar bukan manusia terdeteksi ada puting terjadi pada gambar obyek yang menyerupai puting.

## VI. PENUTUP

### A. Kesimpulan

Perkembangan pornografi dan kemudahan mengaksesnya mengakibatkan kerawanan sosial di kalangan anak di bawah umur. Menggabungkan metode deteksi obyek dan statistik warna kulit dapat dengan tepat mendeteksi gambar pornografi di halaman web. Dengan ini diharapkan ditemukan cara efektif untuk menghindarinya yaitu menemukan aplikasi untuk mengatasi penyebaran pornografi.

1. Dapat mendeteksi puting dengan metode viola jones dengan tingkat akurasi yang baik ini dibuktikan dengan pengujian pada gambar porno yang mengandung puting dan tidak

mengandung puting. Hasil deteksi gambar pornografi pada citra dengan manusia dan puting ada dengan tingkat efektivitas 90% , manusia tidak ada puting efektivitas 94% .

2. Dengan menggabungkan deteksi objek puting dan deteksi warna kulit dapat meningkatkan tingkat akurasi mendeteksi gambar porno.
3. Deteksi dengan viola jones membutuhkan waktu singkat dibanding dengan metode deteksi kulit untuk mendeteksi gambar porno.

#### **B. Saran**

1. Saran untuk penelitian lebih lanjut dan penyempurnaan penelitian tentang deteksi pornografi adalah dengan memperbanyak jumlah sample dan stage classifier .
2. Menambahkan tipe data positif tidak hanya puting.

#### **DAFTAR PUSTAKA**

- [1][3] Rea, Michael C., What is Pornography?, University of Delaware, NOUS 35:1, 118 - 145, Blackwell Publishers Inc., USA, 2001.
- [2] Lembaran Negara Republik Indonesia, Undang-Undang Republik Indonesia Nomor 44 Tahun 2008 Tentang Pornografi, LNRI Nomor 181, Jakarta, 2008
- [4] Zubaedah,Neng, Pengkajian Hukum Tentang Produk Pornografi Untuk Tujuan Dan Kepentingan Pendidikan Dan Layanan Kesehatan , Perpustakaan Mahkamah Agung, Edisi :ISBN/ISSN : 978-602-8815-42-0 ,2012
- [5] [7] Lin ea, Yu-Chun, et.al., Pornography Detection Using Support Vector Machine, 16th IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP 2003): pp 123 – 130, Kinmen, ROC, 17 – 19 Agustus 2003.
- [8] Agbinya, Johnson, et.al., Automatic online porn detection and tracking, 12th IEEE International Conference on Telecommunications (p. 1 – 7), Institute of Electrical and Electronics Engineers (IEEE), Cape Town, South Africa, 2005.
- [9][10]Viola, Paul, dan Jones, Michael, Rapid object detection using a boosted cascade of simple features, Proceedings of the 2001 *IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2001 (CVPR 2001)* (Volume: 1), pp: I-511 – I-518, IEEE, 2001