

## DETEKSI SITUS PORNOGRAFI BERDASARKAN GAMBAR MENGGUNAKAN METODE *CONVOLUTIONAL NEURAL NETWORK*

Ahmad Hunaepi<sup>1</sup>, M. Makhsun<sup>2</sup>, S. Sarwani<sup>3</sup>

<sup>1,2</sup>Program Magister Komputer Program Studi Teknik Informatika  
Sekolah Tinggi Manajemen Informatika dan Komputer Eresha, Jakarta

<sup>3</sup>Pasca Sarjana Magister Manajemen

Universitas Pamulang, Tangerang Selatan Banten

<sup>1</sup>ahmadhunaepi@gmail.com, <sup>2</sup>makhsun\_toha@yahoo.com, <sup>3</sup>sarwani3082@yahoo.com

### ABSTRACT

Today, the development of information technology is very fast. It is also followed by the growth of internet sites that are very fast, no exception the growth of sites contains pornographic. The Association of Indonesian Internet Service Providers (APJII), in 2018, reports as many as 55.9% of Indonesian internet users had experienced of appearance pornographic content suddenly. This study aims to detect pornographic sites based on images using the convolutional neural network method. To find out the level of accuracy in detection of pornographic sites, a training process with a dataset of 300 images is required. From the test results, an accuracy rate of 85% in the detection of pornographic sites is obtained.

**Keywords:** *Detection, Website, Pornography, Convolutional Neural Networks*

### ABSTRAK

Perkembangan teknologi informasi yang begitu cepat diikuti oleh pertumbuhan situs internet yang juga sangat cepat, tidak terkecuali pertumbuhan situs yang bermuatan konten pornografi. Laporan yang dikeluarkan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) pada tahun 2018, sebanyak 55.9% dari pengguna internet Indonesia pernah mengalami muncul konten porno secara tiba-tiba. Penelitian ini bertujuan untuk mendeteksi situs pornografi berdasarkan gambar menggunakan metode *convolutional neural network*. Untuk mengetahui tingkat keakuratan dalam deteksi situs pornografi, diperlukan proses pelatihan dengan *datasets* berjumlah 300 gambar. Dari hasil pengujian, didapatkan nilai tingkat akurasi sebesar 85% dalam deteksi situs pornografi.

**Kata Kunci:** *Deteksi, Website, Pornografi, Convolutional Neural Networks*

DOI: <http://dx.doi.org/10.15408/jti.v12i2.12488>

## I. PENDAHULUAN

### 1.1 Latar Belakang

Internet menjadi jendela informasi di era milenial saat ini. Dengan pesatnya perkembangan internet dan jejaring sosial, ada sejumlah besar gambar yang tersedia bagi pengguna. Sayangnya, banyak situs ilegal yang memposting banyak gambar porno yang dapat menyebabkan masalah sosial yang serius. Misalnya, gambar porno yang dapat dilihat oleh anak-anak di bawah umur. Menurut laporan survei 2018 yang dilakukan oleh Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) [1], sebanyak 55.9% dari pengguna internet Indonesia pernah mengalami muncul konten porno secara tiba-tiba pada konten yang sedang dikunjungi, sedangkan pengguna internet di Indonesia berdasarkan usia pada tahun 2018 terbanyak pada rentang usia 15-19 tahun.

Penelitian yang dilakukan oleh Tri Suyatno [2] mengenai “pengaruh pornografi terhadap perilaku belajar siswa (studi kasus: sekolah menengah x)”, dilakukan survei terhadap siswa SMP-SMA di wilayah Bogor dengan jumlah responden 162 siswa. Kesimpulan dari hasil penelitian terdapat sebanyak 157 dari 162 siswa atau 96.91% siswa pernah menonton pornografi. Sebanyak 59 siswa atau 42.57% menganggap pornografi mengganggu konsentrasi belajar, 36 siswa atau 22.22% menganggap pornografi hanya terkadang mengganggu konsentrasi belajar. Media yang paling banyak digunakan oleh siswa untuk mengakses konten pornografi adalah di warung internet sebanyak 61.11%.

Kementerian Komunikasi dan Informatika (Kominfo) dalam hal ini berperan sebagai pemerintah dalam mengontrol komunikasi data internet di Indonesia, telah berupaya dalam pemblokiran situs-situs pornografi, bahkan hingga tahun 2018 Kominfo telah memblokir situs pornografi lebih dari 800 ribu *website* [3]. Kominfo melakukan pemblokiran dengan cara melakukan *blacklist* terhadap situs-situs porno ke dalam DNS nawala, DNS nawala merupakan layanan penyaringan DNS yang bebas biaya dan dapat digunakan oleh semua pengguna internet [4].

Setiap tahunnya pertumbuhan situs-situs porno terus meningkat, dengan demikian perlu adanya suatu metode baru yang dibutuhkan untuk melakukan pemblokiran terhadap situs-situs pornografi yang semakin meningkat.

Salah satu teknik yang digunakan dalam menangani masalah ini adalah teknik klasifikasi Jaringan Syaraf Tiruan (JST). Jaringan syaraf tiruan merupakan jaringan dari sekelompok unit pemroses kecil yang dimodelkan berdasarkan jaringan saraf yang terdapat pada manusia. JST merupakan sistem adaptif yang dapat mengubah strukturnya untuk memecahkan masalah berdasarkan informasi eksternal maupun internal yang mengalir melalui jaringan tersebut, dengan kata lain teknik ini memiliki kemampuan untuk belajar dari pengalaman. Jenis model jaringan syaraf tiruan yang memiliki beberapa lapisan yang disebut dengan *Multi-Layer Perceptron* (MLP) yang menghubungkan secara penuh antar neuronnya memiliki kemampuan klasifikasi yang *powerful*.

Penelitian-penelitian mengenai *convolutional neural network* sudah pernah dilakukan sebelumnya, diantaranya penelitian yang dilakukan oleh Limantoro, dkk [6] tentang deteksi pengendara sepeda motor. Penelitian ini bertujuan untuk memahami arsitektur CNN dengan hasil yang terbaik. Selain itu penelitian ini juga bertujuan untuk memudahkan pemerintah dalam hal mengawasi kendaraan bermotor yang melanggar aturan lalu lintas. Kesimpulan yang didapatkan dari penelitian ini yaitu penggunaan *Convolution Neural Network* mampu mengklasifikasi gambar yang kompleks seperti pengendara sepeda motor dan bukan pengendara sepeda motor. Hasil dari penelitian memiliki tingkat akurasi sebesar 93%.

Penelitian yang dilakukan oleh Zufar, dkk [7] tentang “*Convolutional Neural Networks* untuk Pengenalan Wajah Secara *Real-Time*”. Penelitian ini menggunakan mengimplementasikan metode CNN dan bantuan *library* OpenCV untuk deteksi multi wajah dan perangkat *Web Cam M-Tech 5MP*. Hasil uji coba dengan menggunakan konstruksi model CNN sampai kedalaman 7 lapisan dengan input dari hasil ekstraksi *Extended Local Binary Pattern* dengan radius 1 dan neighbor 15 memperlihatkan kinerja pengenalan wajah meraih rata-rata tingkat akurasi lebih dari 89% dalam  $\mp 2$  *frame* per detik.

Penelitian yang dilakukan oleh Danukusumo [8] dalam tesisnya dengan judul “Implementasi *Deep Learning* Menggunakan *Convolutional Neural Network* Untuk Klasifikasi Citra Candi Berbasis GPU”. Penelitian ini menggunakan citra candi sebagai

data input. Hasil pengujian menunjukkan akurasi sebesar 98,99% pada *training set* dan 85,57% pada *test set* dengan waktu pelatihan mencapai 389,14 detik. Sehingga dapat disimpulkan bahwa klasifikasi citra candi dapat dilakukan dengan sangat baik menggunakan teknik *Deep Learning* dengan CNN.

Penelitian yang dilakukan oleh Razi [9] dengan judul “Klasifikasi Artikel Berita Berbahasa Indonesia Menggunakan *Convolutional Neural Network*”. Dalam penelitian tersebut digunakan metode *Convolutional Neural Network* (CNN) untuk mengklasifikasikan artikel berita berbahasa Indonesia. Terdapat lima kelas dalam pengklasifikasian, yaitu *entertainment*, kesehatan, olahraga, teknologi, dan ekonomi. Sebelum dilakukan klasifikasi, terlebih dahulu dilakukan perubahan kata-kata ke dalam bentuk vektor dengan menggunakan *word2vec* sehingga hasil dari perubahan tersebut dapat diinputkan ke dalam CNN. Hasil pengujian menunjukkan bahwa kombinasi penggunaan metode *Convolutional Neural Network* dan *word2vec* memberikan hasil nilai akurasi sebesar 96.70 %, dan presisi, *recall* serta *f-measure* mencapai 96,60 %.

Penelitian lain telah dilakukan oleh Rismiyati [10] tentang pengenalan motif batik Pesisir menggunakan salah satu metode *deep learning*, yaitu *Convolutional Neural Network* (CNN). Rismiyati menerapkan algoritma *Convolutional Neural Network* (CNN) untuk membedakan salak yang lolos ekspor dan tidak. *Input* yang digunakan dalam penelitian ini adalah citra salak. Proses yang dilakukan adalah pengumpulan data, proses *preprocessing*, klasifikasi dan pengujian. Klasifikasi digunakan dengan CNN, dimana untuk mendapatkan akurasi terbaik parameter-parameter yang ada harus diujikan. Terdapat dua model yang digunakan yaitu model dua kelas dan empat kelas. Penelitiannya menunjukkan bahwa akurasi terbaik untuk model dua kelas didapatkan metode CNN dengan menggunakan *learning rate* 0.0001, satu lapisan konvolusi dengan jumlah filter lima belas dengan ukuran 3x3x3, dan jumlah neuron pada lapisan tersembunyi 100. Akurasi yang didapatkan adalah 81.5%. Model empat kelas mendapatkan akurasi 70.7% dengan dua lapisan konvolusi.

Berdasarkan penjelasan kelebihan CNN tersebut, dapat diambil kesimpulan bahwa CNN memiliki kemampuan klasifikasi yang

diperuntukan untuk data gambar, sehingga pada penelitian ini akan melakukan deteksi situs porno terhadap gambar menggunakan metode CNN. Gambar porno yang terdapat pada suatu *website* akan dilakukan klasifikasi, sehingga akan didapatkan *website* yang terdapat *content* gambar porno. Diharapkan klasifikasi menggunakan metode CNN mencapai akurasi minimal 80%.

## 1.2 Landasan Teori

### 1.2.1 Citra Digital

Citra merupakan suatu gambar pada bidang dua dimensi. Ditinjau secara matematis citra adalah suatu fungsi *continue* dari intensitas cahaya pada bidang dua dimensi. Sumber cahaya yang menerangi objek akan dipantulkan kembali sebagian berkas cahaya tersebut oleh objek. Pantulan cahaya akan ditangkap oleh alat *optic*, seperti mata pada manusia, pemindai (*scanner*), kamera, dan lain-lain, sehingga citra akan terekam oleh bayangan objek [11].

Citra digital merupakan citra yang diolah secara digital, citra digital diidentifikasi

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & \vdots & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N-1) \end{bmatrix}$$

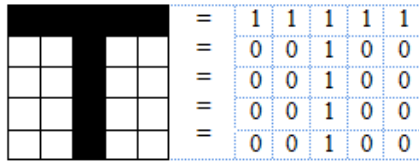
Gambar 1. Gambar matriks citra digital

oleh sebuah matriks dimana baris dan kolom memiliki nilai yang menyatakan suatu titik pada citra. Berikut merupakan gambaran matriks terhadap citra digital:

Nilai pada irisan di antara baris dan kolom (posisi  $x, y$ ) disebut dengan istilah *pixel*. Citra digital terbagi menjadi tiga bagian, yaitu:

#### 1. Citra Biner (Monokrom)

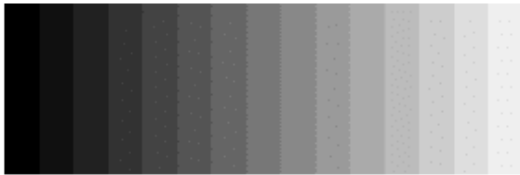
Pada citra ini warna yang dimiliki hanya berjumlah 2, yaitu warna hitam dan putih. Warna hitam memiliki nilai 1 dan warna putih memiliki nilai 0. Kedua warna ini membutuhkan 1bit memori sebagai penyimpanan. Contoh susunan *pixel* pada citra biner sebagai berikut:



Gambar 2. Citra biner

## 2. Citra Grayscale (skala keabuan)

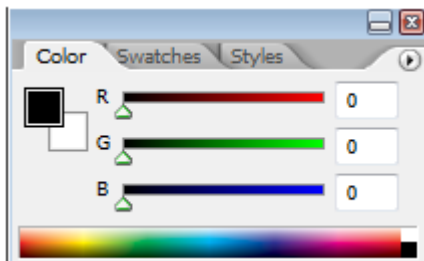
Citra *grayscale* mempunyai kemungkinan warna antara warna hitam sebagai nilai minimal dan warna putih sebagai nilai maksimal. Banyaknya jumlah warna yang terdapat pada citra *grayscale* bergantung pada jumlah bit yang disediakan oleh memori yang akan menampung kebutuhan warna tersebut. Karena semakin banyaknya jumlah bit warna akan berpengaruh dengan semakin halusnya gradasi warna yang akan terbentuk.



Gambar 3. Citra grayscale

## 3. Citra Warna (*true color*)

Setiap citra memiliki ukuran pikselnya masing-masing, pada setiap *pixel* terdapat nilai warna. Nilai warna dari suatu *pixel* merupakan suatu kombinasi dari tiga warna dasar, yaitu warna merah, hijau dan biru (RGB). Setiap warna dasar memiliki penyimpanan 8 bit = 1 *byte*, dalam setiap 1*byte* memiliki nilai dengan nilai maksimum 255, jadi dalam satu *pixel* citra warna memiliki ukuran penyimpan 3 *byte*.

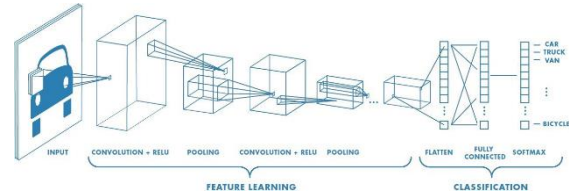


Gambar 4. Citra warna

### 1.2.2 Convolutional Neural Networks

*Convolutional Neural Network* (Jaringan Saraf Tiruan Konvolusi) adalah jenis jaringan saraf tiruan yang terdiri dari beberapa tahap filter dan satu tahap klasifikasi. Tahap filter

digunakan untuk mengekstraksi fitur dari input, yang berisi empat jenis lapisan. Lapisan konvolusional (*convolutional layer*), lapisan aktivasi (*activation layer*), dan lapisan pengumpulan (*pooling layer*). Tahap klasifikasi adalah *perceptron multi-layer*, yang terdiri dari beberapa *layer* yang sepenuhnya terhubung [12]. Cara kerja *Convolutional Neural Network* ditunjukkan pada Gambar 5.



Gambar 5. Convolutional neural network

#### 1.2.2.1. Convolution Layer

*Convolution layer* terdiri dari beberapa neuron yang tersusun, sehingga membentuk filter yang memiliki *pixel* lebar dan tinggi. Filter ini pertama kali memiliki nilai random, dan nilai dari filter pertama yang menjadi parameter yang akan di-*update* pada proses *learning*.

##### 1) *Stride*

*Stride* merupakan parameter yang akan menentukan berapa jumlah pergeseran pada filter. Jika nilai *stride*-nya 2, maka lapisan konvolusi bergeser sebanyak 2 *pixels* yang dimulai secara horisontal lalu vertikal.

Semakin kecil *stride* yang digunakan, maka semakin besar pula informasi yang didapatkan dari sebuah input, namun memerlukan tingkat komputasi yang lebih jika dibandingkan penggunaan *stride* yang besar.

##### 2) *Padding*

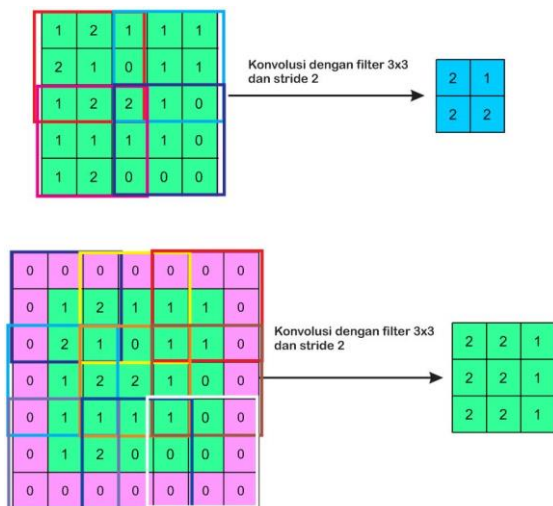
*Padding* atau *Zero Padding* merupakan parameter yang menentukan jumlah *pixels* yang ditambahkan pada setiap sisi dari input. Hal ini bertujuan untuk manipulasi dimensi output dari lapisan konvolusi (*Feature Map*).

Tujuan penggunaan *padding* yaitu:

- Dimensi output dari lapisan konvolusi lebih kecil dari input (kecuali penggunaan 1x1 filter dengan 1 *stride*). Output akan digunakan kembali sebagai input dari lapisan konvolusi selanjutnya, mengakibatkan semakin banyaknya informasi yang terbuang. Dengan menggunakan *padding*, dimensi output dapat diatur agar tetap

sama dengan dimensi input atau tidak berkurang secara drastis. Sehingga kita dapat menggunakan lapisan konvolusi yang lebih dalam agar lebih banyak *features* yang dapat di-*extract*.

- Dapat meningkatkan performa dari model karena lapisan konvolusi fokus pada informasi yang sebenarnya, yang terdapat di antara *zero padding*.



Gambar 6. Penambahan *zero padding*

Pada Gambar 6, dimensi dari input yang sebenarnya yaitu 5x5, jika dilakukan konvolusi menggunakan filter 3x3 dengan *stride* sebesar 2, maka akan didapatkan *feature map* berukuran 2x2. Jika kita menambahkan padding sebanyak 1, maka akan didapatkan *feature map* yang berukuran 3x3 lebih banyak informasi yang didapatkan.

Untuk menghitung dimensi dari *feature map*, dapat menggunakan rumus seperti di bawah ini:

$$\text{Output} = \frac{(W - N + 2P)}{S} + 1 \tag{1}$$

- $W$  = Panjang/Tinggi Input
- $N$  = Panjang/Tinggi Filter
- $P$  = *Zero Padding*
- $S$  = *Stride*

### 2.2.2. ReLU Layer

*ReLU Layer* merupakan fungsi yang paling umum digunakan untuk output dari lapisan konvolusi untuk pemetaan *nonlinear*. Dalam matematika, ReLU dapat dinyatakan sebagai berikut:

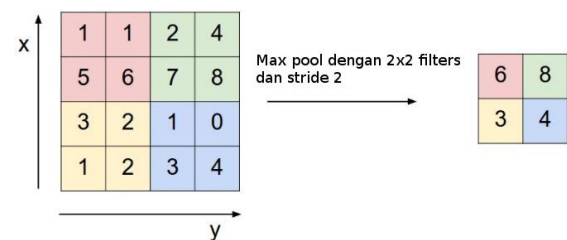
$$f(x) = \max(0, x) \tag{2}$$

Untuk jaringan konvolusi, ReLU lebih efektif daripada fungsi sigmoid dan lebih praktis daripada tangen hiperbolik. Karena alasan ini, ReLU telah menjadi fungsi aktivasi paling populer untuk CNN.

### 2.2.3. Pooling Layer

Letak *pooling layer* biasanya terletak setelah lapisan konvolusi, *pooling layer* merupakan sebuah filter dengan ukuran dan *stride* tertentu, biasanya bekerja dengan cara bergeser pada seluruh area *feature map*.

Yang biasa digunakan pada *pooling layer* yaitu *max pooling* dan *average pooling*. Sebagai contoh *max pooling* bisa dilihat pada Gambar 7.

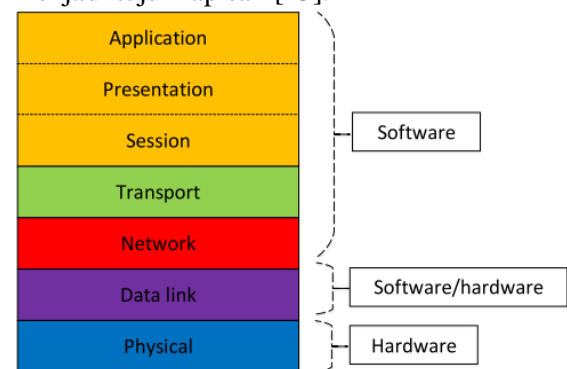


Gambar 7. *Max Pooling 2x2*

Pada penggunaan *max pooling 2x2* dengan *stride 2*, maka setiap pergeseran filter pada area 2x2 pixel akan dipilih nilai yang paling besar. Berbeda dengan *average pooling*, nilai yang akan keluar yaitu nilai rata-rata dari setiap pergeseran filter.

### 1.2.3 Open System Interconnection

*Open System Interconnection (OSI)* berfungsi sebagai alat referensi untuk komunikasi antara perangkat yang terhubung dalam suatu jaringan. OSI adalah suatu kerangka kerja konseptual yang membantu kita memahami interaksi yang kompleks dalam jaringan. OSI membagi proses komunikasi menjadi tujuh lapisan [13].



Gambar 8. *OSI layers communication*

Setiap lapisan memiliki fungsi dan saling terhubung antara lapisan atas dan bawah. *Application layer* merupakan lapisan atas dari model OSI yang memiliki fungsi untuk menyediakan sebuah antarmuka antara aplikasi dengan protokol jaringan; *presentation layer* berfungsi untuk memastikan bahwa data berada dalam format yang dapat digunakan dan dapat melakukan proses enkripsi/deskripsi informasi; *Session layer* berfungsi untuk mengontrol *port* dan sesi, seperti menetapkan, dan memutuskan koneksi antar komputer; *Transport layer* berfungsi untuk mentransmisikan data menggunakan protokol transmisi termasuk TCP dan UDP; *Network layer* berfungsi untuk menetapkan jalur yang digunakan untuk melakukan transfer data, *network layer* juga dapat membaca *IP address* dari paket data; *Data link layer* berfungsi untuk membaca *MAC address* dari paket data dan bertanggung jawab memeriksa kesalahan yang akan terjadi pada saat transmisi data; *physical layer* berfungsi untuk mengirimkan data melalui media fisik.

#### 1.2.4 Website

*Website* merupakan salah satu aplikasi yang berisikan kumpulan teks, gambar, suara, animasi, video, dokumen di dalamnya. *Website* berjalan pada protokol HTTP (*hyper text transfer protocol*), untuk mengakses *website* bisa menggunakan browser. Terdapat beberapa jenis *browser* populer saat ini, di antaranya: Google Chrome, Firefox, Safari, dan lain-lain [14].

#### 1.2.5 Wireshark

*Wireshark* merupakan suatu aplikasi penganalisa yang paling populer. *Wireshark* memiliki kemampuan untuk memindai segala jenis jaringan *ethernets*, Wi-Fi, bahkan hingga *bluetooth*. Ini adalah alasan utama mengapa *wireshark* digunakan secara global dalam raksasa industri tidak hanya sebagai pelacak tetapi juga sebagai *Intrusion Detection System* (IDS). Aplikasi ini dapat dengan cepat mendeteksi dan melaporkan serangan *Denial of Service* (DOS), serangan pada *firewall* atau pelanggaran keamanan apapun melalui jaringan. *Wireshark* dapat mencapai ke dalam bit paket, dan juga dapat memisahkan paket protokol jaringan yang berbeda dengan menyoroti protokol dengan skema warna yang berbeda seperti hijau untuk HTTP, biru untuk DNS, dan lain-lain [15].

## II. METODOLOGI

### 2.1 Metode Pemilihan Data

Untuk melakukan penelitian dengan pendekatan kuantitatif dibutuhkan sumber daya yang besar dalam hal waktu dan uang, hal ini sulit dipenuhi oleh seorang mahasiswa, sehingga menganalisa data sekunder merupakan strategi yang lebih bermanfaat [16]. Maka pada penelitian ini digunakan data sekunder.

### 2.2 Metode Pengumpulan Data

Beberapa publikasi penelitian menggunakan *dataset* yang tersedia secara umum di internet. Pada penelitian ini juga menggunakan *dataset* yang tersedia dari Internet, sehingga metode pengumpulan data yang digunakan adalah dengan mengunduh *dataset porn data* dari <https://www.kaggle.com/ljlr34449/porn-data>. Spesifikasi dan atribut *dataset porn* data yang diperoleh ditunjukkan pada tabel berikut.

Tabel 1. Spesifikasi dan atribut dataset

No	Atribut	Contoh
1	<i>img_source</i>	<a href="http://cdn-d-img.pornhub.com/m=eafT8daaaa/videos/201010/28/84149/original/12.jpg">http://cdn-d-img.pornhub.com/m=eafT8daaaa/videos/201010/28/84149/original/12.jpg</a>
2	<i>length</i>	1583.0
3	<i>nb_views</i>	127450.0
4	<i>quality</i>	LOW
5	<i>title</i>	Tease then suck
6	<i>video_link</i>	<a href="http://thumbzilla.com/video/244ff4e3d92c3bc9433b/tease-then-suck">http://thumbzilla.com/video/244ff4e3d92c3bc9433b/tease-then-suck</a>
7	<i>voting</i>	84.0
8	<i>categories</i>	Blowjob
9	<i>tags</i>	blowjob__oral__fullvideo__strip__tease

### 2.3 Perancangan Penelitian

Di bawah ini merupakan alur yang penulis lakukan berdasarkan kerangka pemikiran di



atas. Berikut penjelasan dari kerangka pemikiran di atas:

- a. Identifikasi Masalah  
Di tahap identifikasi masalah, penulis melakukan identifikasi masalah yang ada pada sistem yang sedang berjalan agar dapat menetapkan masalah yang terjadi serta dapat mengambil keputusan untuk solusi terbaik.
- b. Pustaka Literatur  
Pada tahap ini penulis mengumpulkan beberapa jurnal sebagai bahan acuan yang dapat mendukung dalam proses penelitian ini.
- c. Pengumpulan Data  
Setelah mengidentifikasi masalah yang ada dan telah membaca beberapa pustaka literatur. Penulis melakukan pengumpulan data yang dibutuhkan guna mendukung proses penelitian ini.
- d. Pengelompokan Data  
Pada tahap ini seluruh data yang telah dikumpulkan pada tahap sebelumnya akan dikelompokkan berdasarkan kategori data guna mempermudah dalam proses pengolahan data.
- e. Pengolahan Data  
Tahap akhir adalah penerapan data yang akan dilakukan proses training dan testing menggunakan *Convolutional Neural Network* untuk mendeteksi situs berdasarkan konten negatif.

## 2.4 Teknik Analisis

Dalam penelitian ini, setelah data didapatkan ada beberapa tahapan yang dilakukan terhadap data yang diperoleh. Tahapan tersebut antara lain:

1. Melakukan pengelompokan data gambar sesuai dengan kategori, penulis membuat 2 kelas kategori yaitu gambar porno dan gambar bukan porno.
2. Menerapkan metode *Convolutional Neural Network* sebagai proses pelatihan untuk mendapatkan performa akurasi terbaik. Dalam penelitian ini, penulis merancang dengan kedalaman 7 layer model konvolusi antara lain yaitu *input layer*, *convolutional layer C1*, *pooling layer P2*, *convolutional layer C3*, *pooling layer P4*, *hidden layer H5* dan *output layer O6*.
3. Melakukan pengujian terhadap hasil pelatihan yang dilakukan sebelumnya, dengan cara memberikan beberapa gambar yang akan diklasifikasikan, sehingga dapat

ditemukan nilai akurasi dari proses pengujian, untuk menentukan nilai akurasi menggunakan rumus:

$$Akurasi = \left( \frac{\text{jumlah data yang benar}}{\text{jumlah data keseluruhan}} \right) \times 100\% \quad (3)$$

4. Melakukan pengujian terhadap situs porno, dengan melakukan ekstraksi gambar dari *packet capture data* yang dibuat dengan aplikasi wireshark, setelah gambar didapatkan akan dilakukan deteksi dengan model yang telah dilatih dan akan didapatkan situs yang terindikasi situs porno.

## III. HASIL DAN PEMBAHASAN

### 3.1 Pengumpulan Datasets

Untuk proses melakukan deteksi situs porno berdasarkan gambar dibutuhkan *datasets* berupa gambar yang dilakukan untuk proses pelatihan dan pengujian. Data berupa gambar porno didapatkan pada situs kaggle.com yang merupakan *website* dengan kumpulan-kumpulan *datasets*. Dan data berupa gambar bukan porno didapatkan dengan cara melakukan pengumpulan data pada google gambar dengan menggunakan *software fatkun batch image* yang dapat melakukan pengunduhan gambar secara banyak. Gambar yang digunakan berjumlah 380 gambar yang dibagi menjadi 3 bagian, yaitu 300 gambar untuk dilakukan proses pelatihan, 50 gambar untuk dilakukan proses validasi, dan 30 gambar untuk dilakukan proses pengujian. Contoh beberapa gambar porno dan bukan porno dapat dilihat pada Gambar 10.

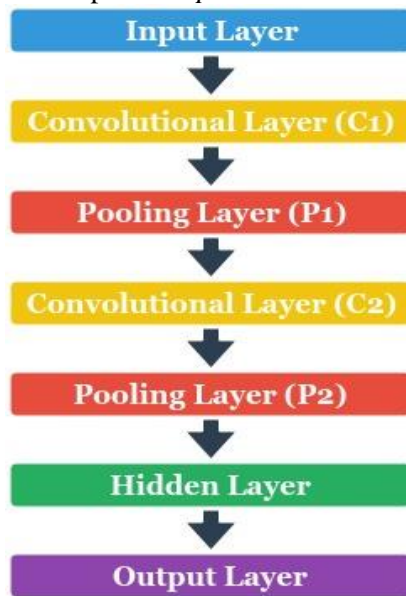


Gambar 9 Datasets Gambar

### 3.2 Proses Training

Pada proses *training* menggunakan metode *Convolutional Neural Networks* ini digunakan 7 *neuron layer* yaitu lapisan *input*, lapisan

konvolusi C1, lapisan *pooling* P1, lapisan konvolusi C2, lapisan *pooling* P2, lapisan *hidden* dan lapisan *output*.

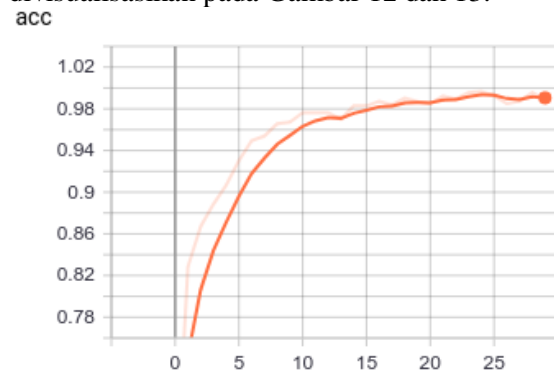


Gambar 10. Arsitektur CNN pelatihan

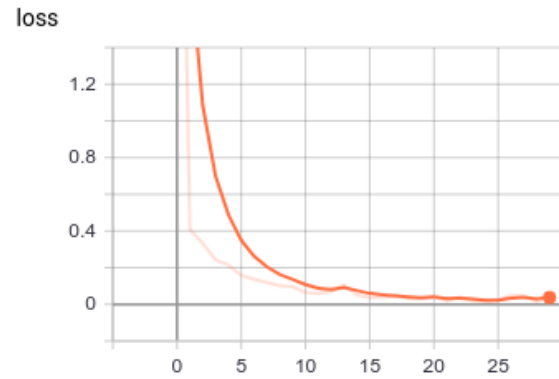
Adapun aplikasi yang digunakan dalam proses pelatihan ini, menggunakan aplikasi *Python* dengan *framework tensorflow* dan *keras*.

### 3.3 Analisis Model Training

Hasil dari percobaan pada proses *training* menggunakan CNN mendapatkan akurasi dari bobot yang sudah diperbaharui selama 30 *epoch*. Hasil dari percobaan selama 30 *epoch* divisualisasikan pada Gambar 12 dan 13.



Gambar 11. Peforma akurasi *training* CNN



Gambar 12. Tingkat *error training* CNN

Pada Gambar 12 menunjukkan performa nilai akurasi yang dilakukan selama 30 *epoch*, nilai akurasi di *epoch* ke-30 yaitu 96.27% dengan tingkat *error* sebesar 3.73%.

### 3.4 Analisis Uji Coba Pengujian Gambar

Sebelum dilakukan pengujian terhadap situs porno, dilakukan terlebih dahulu pengujian terhadap beberapa gambar. Terdapat 30 gambar dengan komposisi 15 gambar porno dan 15 gambar bukan porno. Pengujian ini dilakukan untuk menguji tingkat keberhasilan metode CNN dalam melakukan klasifikasi. Berikut hasil dari beberapa pengujian.

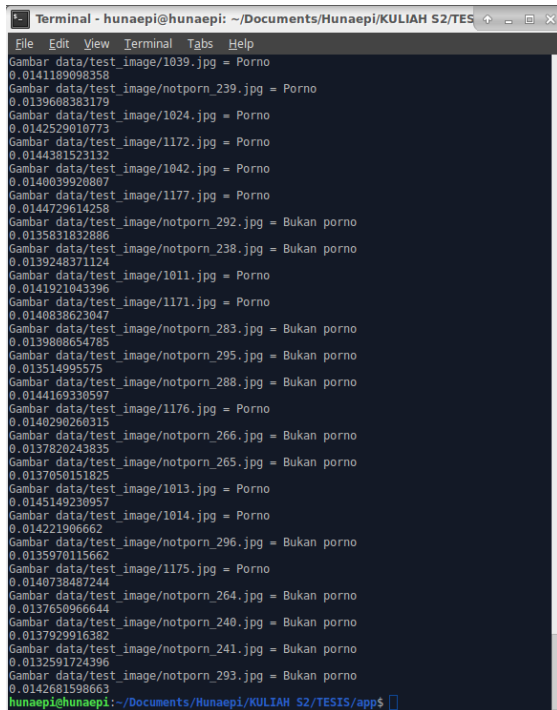
Tabel 2. Hasil pengujian citra

Citra Uji	Hasil	Waktu (detik)
1173.jpg	B	0.0472
1012.jpg	B	0.0145
1041.jpg	B	0.0136
notporn_291.jpg	B	0.0142
notporn_239.jpg	S	0.0138
notporn_292.jpg	B	0.0136
Rata-rata		0.0151

Keterangan: B = Benar, S = Salah

- Waktu rata-rata yang dilakukan untuk melakukan deteksi citra *training* yaitu 0.0151.
- Akurasi = (jumlah data benar / jumlah data keseluruhan) x 100% = (29/30) x 100% = 96.6%.
- *Error* = (jumlah data salah / jumlah data keseluruhan) x 100% = (1/30) x 100% = 3.3%.





Gambar 13. Hasil pengujian citra

### 3.5 Datasets Uji Coba Pengujian Situs Porno

Dataset uji coba pengujian dikumpulkan dengan cara melakukan *capture* pada *packet* data jaringan yang telah mengakses beberapa situs menggunakan aplikasi *wireshark*. Adapun situs-situs yang dilakukan *capture* telah dipersiapkan terlebih dahulu dengan membuat tampilan *website* menggunakan HTML yang didalamnya berisi gambar, tampilan *website* masing-masing akan diberikan nama domain yang berbeda-beda. Rincian daftar *datasets* disajikan pada tabel berikut.

Tabel 3. Daftar situs

Domain	Keterangan
siteporn1.local	P
siteporn2.local	P
siteporn3.local	P
siteporn4.local	P
siteporn5.local	P
siteporn6.local	P
siteporn7.local	P
siteporn8.local	P

siteporn9.local	P
siteporn10.local	P
sitenotporn1.local	NP
sitenotporn2.local	NP
sitenotporn3.local	NP
sitenotporn4.local	NP
sitenotporn5.local	NP
sitenotporn6.local	NP
sitenotporn7.local	NP
sitenotporn8.local	NP
sitenotporn9.local	NP
sitenotporn10.local	NP

P = Situs porno  
NP = Situs bukan porno

#### 3.5.1 Proses Input Pengujian Situs Porno

Setelah menentukan beberapa situs porno yang akan diakses. Selanjutnya satu per satu situs dilakukan proses pengujian dengan cara mengakses situs melalui *browser* dan dipastikan aplikasi *wireshark* telah berjalan untuk menangkap proses alur paket data pada situs yang kita akses.



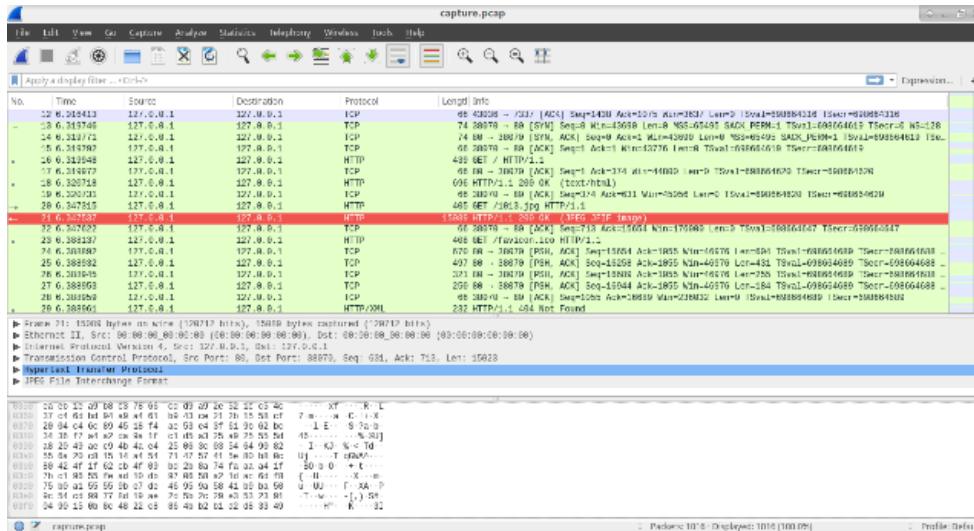
Gambar 14. Contoh situs porno



Gambar 15. Contoh situs bukan porno

#### 3.5.2 Proses Deteksi Situs

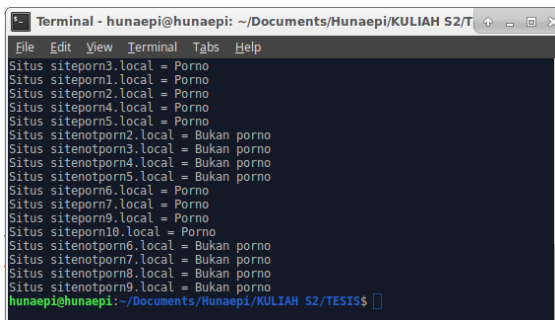
Setelah melakukan proses input pengujian dengan cara mengakses situs melalui *browser*, akan didapatkan data gambar yang terekam oleh aplikasi *wireshark*.



Gambar 16. Packet capture wireshark

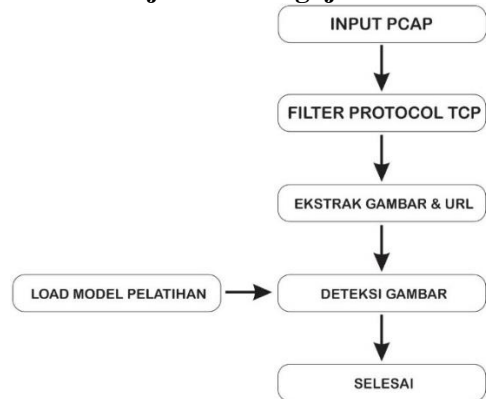
3.5.3 Proses Ekstraksi dan Klasifikasi Situs

Proses ekstraksi dan klasifikasi dilakukan dengan membuat script menggunakan bahasa Python. Penulis menggunakan library scrapy untuk membaca dan mengekstrak gambar melalui file capture yang dibuat pada proses deteksi menggunakan aplikasi wireshark, dan hasil ekstrak gambar tersebut dilakukan klasifikasi menggunakan library tensorflow dan keras.



Gambar 17. Hasil pengujian situs

3.6 Analisis Uji Coba Pengujian Situs Porno



Gambar 18. Proses pengujian

Uji coba pengujian model dilakukan dengan menggunakan packet capture yang sebelumnya telah dibuat menggunakan aplikasi wireshark. Proses deteksi dijalankan dengan cara mengekstrak gambar dari data packet capture, dan dilakukan proses deteksi gambar menggunakan metode convolutional neural network. Berikut daftar hasil dari pengujian model.

Tabel 4. Hasil pengujian

Domain	Hasil
siteporn3.local	B
siteporn1.local	B
siteporn2.local	B
siteporn4.local	B
siteporn5.local	B
sitenotporn2.local	B

sitenotporn3.local	B
sitenotporn4.local	B
sitenotporn5.local	B
siteporn6.local	B
siteporn7.local	B
siteporn9.local	B
siteporn10.local	B
sitenotporn6.local	B
sitenotporn7.local	B
sitenotporn8.local	B
sitenotporn9.local	B
siteporn8.local	S
sitenotporn1.local	S
sitenotporn10.local	S

B = Benar, S = Salah

- Data *website* keseluruhan untuk pengujian sejumlah 20.
- Akurasi = (jumlah data benar / jumlah data keseluruhan) x 100% = (17/20) x 100% = 85%.
- *Error* = (jumlah data salah / jumlah data keseluruhan) x 100% = (3/20) x 100% = 15%.

#### IV. PENUTUP

Dari hasil pembahasan pada penelitian ini dapat disimpulkan yaitu penggunaan *convolutional neural network* dengan kedalaman *7layer* konvolusi dengan akurasi dari hasil *training* sebesar 96.27% yang dilakukan selama 30 *epoch* mampu mengidentifikasi situs pornografi berdasarkan gambar dengan nilai akurasi 85%.

Untuk saran dalam pengembangan penelitian selanjutnya perlu adanya penambahan jumlah sampel dengan 3 klasifikasi yaitu gambar porno, semi porno dan bukan porno. Selanjutnya, perlu adanya implementasi ke dalam sistem, untuk proses filterasi terhadap situs pornografi.

#### DAFTAR PUSTAKA

[1] APJII, “Infografis Penetrasi & Perilaku Pengguna Internet Indonesia,” *Teknopreuner*, vol. 1, no. Hasil Survei

Penetrasi dan Perilaku Pengguna Internet Indonesia 2017, pp. 1–39, 2019.

[2] T. Suyatno, “Pengaruh Pornografi Terhadap Perilaku Belajar Siswa ( Studi Kasus : Sekolah Menengah X ),” *J. Pendidik. Dompot Dhuafa*, pp. 1–12, 2011.

[3] A. T. Haryanto, “Kominfo Sudah Blokir 106.466 Situs Pornografi Tahun Ini,” *www.detik.com*, 2018. [Online]. Available: <https://inet.detik.com/law-and-policy/d-4354570/kominfo-sudah-blokir-106466-situs-pornografi-tahun-ini>. [Accessed: 16-Jul-2019].

[4] I. Rahmayani, “Begini Cara Kominfo Blokir Situs Terlarang,” *kominfo.go.id*, 2015. [Online]. Available: [https://kominfo.go.id/content/detail/5957/begini-cara-kominfo-blokir-situs-terlarang/0/sorotan\\_media](https://kominfo.go.id/content/detail/5957/begini-cara-kominfo-blokir-situs-terlarang/0/sorotan_media). [Accessed: 16-Jul-2019].

[5] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[6] S. E. Limantoro, Y. Kristian, and D. D. Purwanto, “Deteksi Pengendara Sepeda Motor Menggunakan Deep Convolutional Neural Networks,” *Semin. Nas. Teknol. Inf. dan Komun.*, pp. 79–86, 2017.

[7] M. & B. S. Zufar, “Convolutional Neural Networks untuk Pengenalan Wajah Secara Real-Time,” *J. Sains dan Seni*, vol. 5, no. 3, pp. 1–6, 2016.

[8] K. P. Danukusumo, “Implementasi Deep Learning Menggunakan Convolutional Neural Network Untuk Klasifikasi Citra Candi Berbasis GPU,” vol. 4, pp. 9–15, 2017.

[9] A. RAZI, “Klasifikasi Artikel Berita Berbahasa Indonesia Menggunakan Convolutional Neural Network,” 2017.

[10] RISMİYATI, “Implementasi Convolutional Neural Network Untuk Sortasi Mutu Salak Ekspor Berbasis Citra Digital,” 2016.

[11] A. Budi, S. Suma’inna, and H. Maulana, “Pengenalan Citra Wajah Sebagai Identifier Menggunakan Metode Principal Component Analysis (PCA),” *J. Tek. Inform.*, vol. 9, no. 2, pp. 166–175, 2018.

[12] W. Zhang, C. Li, G. Peng, Y. Chen, and

- Z. Zhang, "A deep convolutional neural network with new training methods for bearing fault diagnosis under noisy environment and different working load," *Mech. Syst. Signal Process.*, vol. 100, pp. 439–453, 2018.
- [13] W. Edmonson et al., "Systems engineering of inter-satellite communications for distributed systems of small satellites," 9th Annu. IEEE Int. Syst. Conf. SysCon 2015 - Proc., pp. 705–710, 2015.
- [14] M. Suhartanto, "Pembuatan Website Sekolah Menengah Pertama Negeri 3 Delanggu Dengan Menggunakan Php Dan Mysql," *J. Speed*, vol. 4, no. 1, pp. 1–8, 2012.
- [15] P. Goyal and A. Goyal, "Comparative study of two most popular packet sniffing tools-Tcpdump and Wireshark," in 2017 9th International Conference on Computational Intelligence and Communication Networks (CICN), 2017, pp. 77–81.
- [16] M. Berndtsson, J. Hansson, B. Olsson, and B. Lundell, *Thesis Projects: A Guide for Students in Computer Science and Information Systems*, 2008.