Prosiding Seminar Nasional Universitas Ma Chung
pp. 61-70, 2021
ISSN 2808-7771 online

Deteksi Objek pada Film Menggunakan Yolo Object Detector dan K-Nearest Neighbor

Windra Swastika[1*], Marcellino Agustinus Sinaga [1]

[1]Program Studi Teknik Informatika, Universitas Ma Chung, Jl. Villa Puncak Tidar N-1, Malang, Indonesia, 65142

Correspondence: Windra Swastika (windra.swastika@machung.ac.id)
Received: *15 Agustus 2021* – Revised: *1 November 2021* - Accepted: *3 November 2021*

Abstrak. Peran film sebagai industri budaya merupakan salah satu faktor yang dianggap penting bagi Indonesia. Film harus benar–benar diperhatikan dan dilindungi agar tidak berbalik menjadi pengaruh negatif yang tidak sesuai dan mengakibatkan kemunduran bagi negara Indonesia. Rating usia dari sebuah film merupakan solusi untuk masalah tersebut. Rating usia menunjukkan sentimen pada film tersebut negatif atau positif terutama untuk anak-anak. Pada penelitian ini akan dibuat sebuah sistem untuk mendeteksi objek pada film menggunakan You Only Look Once (YOLO). YOLO digunakan untuk mendapatkan data dari objek yang terdeteksi pada film yaitu untuk objek pistol, pisau, dan rokok dengan jumlah iterasi 50.000. Hasil dari YOLO Mean Average Precision (mAP) adalah 67,14%.

Kata kunci: *K-Nearest Neighbour, object detection, sentiment analysis, YOLO*

Citation Format: Swastika, W. & Sinaga, M.A. (2021). Deteksi Objek pada Film Menggunakan Yolo Object Detector dan K-Nearest Neighbor. *Prosiding Seminar Nasional Sistem Informasi dan Teknik Informatika (SAMA SISI), 61-70*

ISSN 2808-7771 online

PENDAHULUAN

Pentingnya dunia perfilman di Indonesia membuat film harus benar–benar diperhatikan dan dilindungi agar tidak berbalik menjadi pengaruh negatif yang tidak sesuai dan mengakibatkan kemunduran bagi negara Indonesia. Kalangan yang memang harus sangat diperhatikan agar tidak terkena dampak negatif dari sebuah film adalah anak–anak. Batasan umur yang dibuat pada suatu film ditujukan agar anak–anak yang masih dibawah umur agar tidak menonton film yang memiliki konten seperti kekerasan atau konten dewasa [Komalawati, 2017]. Hal itu dilakukan demi menjaga perkembangan mental dan sentimen anak agar tidak terpengaruh hal–hal negatif yang mungkin ditunjukkan dalam film yang dilihat. Penentuan sentimen pada film yang masih dengan melihat keseluruhan film secara manual, dapat ditentukan dengan teknologi yang telah ada di zaman modern ini. Teknologi yang semakin hari semakin maju dengan cepat, terlebih dengan kemunculan AI atau *Artificial Intelligence* membuat berbagai persoalan dapat dipecahkan menggunakan mesin dan teknologi (Aggarwal, 2015).

Salah satu teknologi terbaru saat ini yang dapat digunakan untuk mengambil data pada sebuah video yang dapat digunakan untuk menentukan sentimen pada suatu film adalah *object detector* atau deteksi objek. Teknologi deteksi objek yang mulai berkembang memunculkan banyak metode yang mempunyai keunggulan masing–masing. Salah satu teknologi deteksi objek yang unggul dalam hal kecepatan mengelola gambar adalah *You Only Look Once* atau YOLO (Du, 2018 dan Rosebrock, 2017). Dengan adanya teknologi tersebut, benda–benda yang termasuk dalam kategori kekerasan dan tidak baik sesuai dengan aturan tertulis Komisi Penyiaran Indonesia (KPI) yang mungkin berdampak negatif dapat dideteksi di film tersebut. Hasil deteksi tersebut dapat berguna untuk menjadi data yang dapat diolah untuk menentukan sentimen sebuah film.

METODE PELAKSANAAN

Tahap sebelum membuat dan mengembangkan sistem, persiapan awal perlu dilakukan terlebih dahulu. Model dalam membuat dan mengembangkan sistem dibuat sesuai dan urut sesuai dengan kebutuhan penelitian. Berikut merupakan tahapan yang akan dilakukan pada penelitian ini.

A. Analisis Kondisi

Pada analisis kondisi, dilakukan analisis pada permasalahan yang muncul pada penentuan sentimen film. Analisis dilakukan dengan memperhatikan penentuan sentimen pada film yang saat ini masih menggunakan cara manual dengan melihat film secara

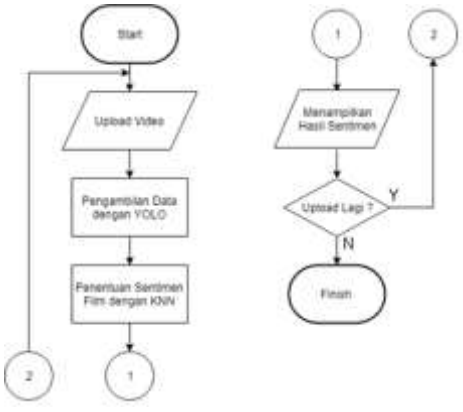
ISSN 2808-7771 online

keseluruhan yang dapat memakan waktu yang cukup lama. Sementara dengan teknologi yang semakin berkembang, hal tersebut dapat ditangani oleh sebuah komputer yang sudah diprogram. Dalam tahapan ini juga dilakukan pengumpulan data berupa film-film yang akan diolah menjadi data-data baru yang akan digunakan untuk penentuan sentimen. Data-data berupa gambar-gambar pistol, rokok, dan juga senjata tajam (pisau) juga akan dikumpulkan dalam tahap ini untuk digunakan sebagai data latih model.

Dataset gambar yang digunakan didapat dari 3 sumber yaitu dataset gambar dari Google *Image, website Soft Computing and Intelligent Information System* dari Universitas Granda dan *Image Collection of Annotated Real-World Objects* (ICARO) 2010. Data yang berhasil diambil dengan pengambilan data sekitar 5000 dan dengan dilakukan *data cleaning* data berkurang menjadi 4323 data gambar (2972 pistol, 861 pisau, dan 490 rokok).

B. Desain Sistem

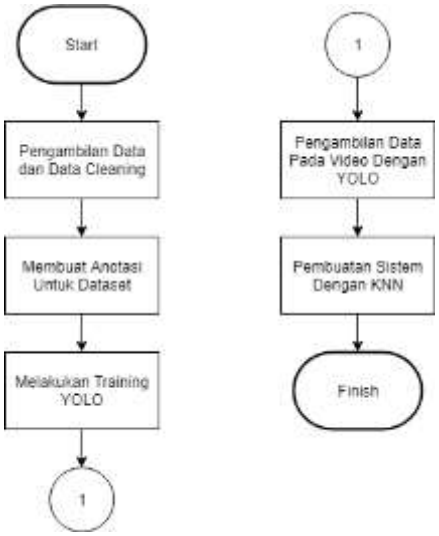
Alur kerja dari sistem penentuan sentimen pada film yang akan dibuat akan digambarkan dalam Flowchart sistem yang dapat dilihat pada Gambar 1. Pada tahap awal *user* dapat melakukan *upload* video yang ingin mereka tentukan sentimennya. Setelah video ter*upload* maka YOLO akan bekerja untuk mendeteksi objek-objek yang negatif pada video tersebut dan mendapatkan datanya. Selanjutnya dari data yang sudah didapatkan oleh YOLO, dapat diproses untuk ditentukan sentimennya menggunakan metode data mining seperti K-Nearest Neighbor atau yang lain [Larose, 2005]. Setelah sentimen sudah didapatkan hasilnya maka hasil tersebut akan ditampilkan kepada *user*.



Gambar 1. Alur kerja sistem

ISSN 2808-7771 online

Sementara untuk proses pembuatan sistem dapat dilihat pada Flowchart Gambar 2. Tahap awal adalah pengambilan data beserta dengan *data cleaning*. Setelah semua data sudah terkumpul dengan baik, maka setiap data perlu diberikan anotasi sebelum masuk ke proses *training*. Jika dataset gambar beserta anotasinya sudah siap maka masuk ke tahap *training* model YOLO. Saat *training* sudah terselesaikan maka model dapat digunakan untuk mengambil data pada data video atau film yang sudah dikumpulkan. Data dari deteksi objek YOLO tersebut yang akan digunakan untuk menentukan sentimen film menggunakan KNN.



Gambar 2. Alur kerja perancangan sistem

Pembuatan sistem dibuat dengan beberapa bahasa pemrograman yang mendukung pembuatan sistem ini. Pada tahap implementasi sistem ke dalam sebuah baris *coding*, banyak hal yang perlu diperhatikan dan dilakukan demi mendapatkan hasil yang baik. Perbaikan akan terus dilakukan demi mendapatkan hasil yang lebih baik sebelum melanjutkan tahap pengujian.

Sistem ini akan dibuat dengan berbagai *software* dan bahasa pemrograman sebagai berikut.

a. Data : *File CSV*
b. Pemrograman : Python 3.5.4 dan C
c. *Framework* : Flask
d. Deteksi Objek : YOLO v3
e. *Browser* : Google Chrome

Selain itu *hardware* yang dipakai untuk membuat dan mengembangkan sistem ini adalah sebagai berikut.

ISSN 2808-7771 online

a. Sistem Operasi : Windows 10 Pro
b. Prosesor : AMD Rayzen1 400
c. GPU : Nvdia GTX 1050 TI
d. RAM : 8GB

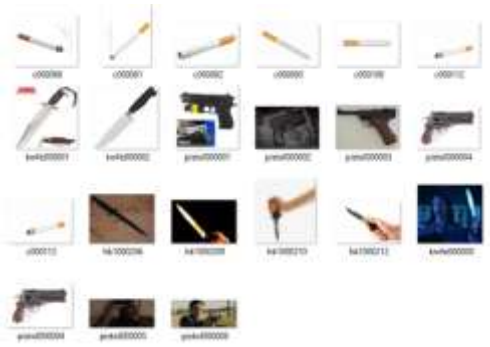
Sedangkan untuk pengujian sistem akan dilakukan beberapa tahap, yaitu:

a. Menguji akurasi deteksi objek YOLO dengan *testing image* dari 10% *dataset* yang sudah didapatkan dan menghitung presisi, *recall*, dan *Mean Average Precision* (mAP) dari model.
b. Menguji akurasi deteksi objek dengan *frame-frame* yang terdeteksi pada film dengan menghitung akurasi, sensitivitas, dan spesifitasnya.

HASIL DAN PEMBAHASAN

Hasil Pengumpulan Data

Data yang diambil berupa data gambar yang akan digunakan sebagai *dataset* untuk pembuatan model deteksi objek YOLO. Sebanyak total 5400 data gambar yang dapat dilihat pada Gambar 3 berhasil didapatkan dan juga terdapat 3600 data anotasi berupa *file txt* yang berisi nama *label* objek beserta dengan *bounding box* atau letak dari objek dalam gambar yang berhasil diambil dari ketiga sumber tersebut.



Gambar 3. Contoh dataset gambar

Selanjutnya adalah mengelolah *dataset* dengan proses *data cleaning* yang berguna untuk membuang data-data yang tidak relevan terhadap objek yang digunakan dan juga objek-objek yang tidak jelas atau ambigu. Proses *data cleaning* juga ditujukan untuk meningkatkan keakuratan data latih dan data uji atau dapat disebut sebagai pengoptimalan data untuk model yang akan dibuat.

ISSN 2808-7771 online

Model YOLO V3

Ada beberapa tahap yang harus dilakukan untuk pembuatan model YOLO V3 mulai dari konfigurasi model sesuai dengan jumlah objek sampai dengan pelatihan model hingga model dapat digunakan untuk mendeteksi objek.

Konfigurasi model yang ada pada YOLO dilakukan dari pengaturan *file model.cfg* yang berisi berbagai macam parameter masukkan untuk model YOLO yang digunakan seperti jumlah konvolusi, *filter*, *strides*, dan jumlah kelas atau objek. Pada penelitian ini yang menggunakan 3 jenis objek, konfigurasi yang harus diganti dalam pengaturan *model.cfg* adalah jumlah kelas yang digunakan, *anchors*, dan juga *filter* sesuai dengan jumlah objek. Perubahan jumlah kelas yang digunakan pada model diubah pada baris 135 dan 177. Sementara nilai untuk perubahan *anchors* dapat didapatkan melalui program *darknet.exe* dengan argumen *calc_anchors*. Nilai *anchors* yang didapatkan digunakan sebagai pengganti *anchors* pada konfigurasi model di baris ke 134 dan 176. Lalu jumlah *filter* yang digunakan diubah sesuai dengan jumlah jenis objek. Hasil yang didapatkan digunakan untuk mengubah nilai pada baris pengaturan ke 127 dan 171.

Setelah perubahan pengaturan *model.cfg* selesai dilakukan, maka masuk ke dalam pembuatan *file* yang diperlukan untuk model YOLO yaitu *obj.names* dan *obj.data* [Redmon, 2016]. Pada *file obj.names* berguna untuk memberi nama objek dan juga mengurutkan jenis objeknya. Sementara pada *file obj.data* berguna untuk mengatur data yang akan dipakai untuk model di mana data yang dipakai untuk *training* atau *testing* yang diatur pada *file train.txt* dan *test.txt*, data yang dipakai sebagai nama objek, jumlah kelas, dan tempat penyimpanan bobot model yang sudah dilatih. Setelah *obj.data* berhasil dibuat, tahap selanjutnya adalah membuat *file train.txt* dan *test.txt* yang berguna untuk memisah data yang digunakan untuk data latih dan data uji. Sementara isi didalam kedua *file* tersebut adalah *path* dari *dataset* yang sudah disiapkan.

Hasil Pengambilan Data Film

Model dari YOLO yang sudah dilatih sebelumnya sudah dapat digunakan untuk mengambil data-data dari ketiga objek yang ada dari sebuah video film. Data-data video film yang akan digunakan diambil dan didapatkan dari media *online* atau internet. Jumlah data film yang diambil berjumlah 24 film dan masing-masing sentimen (positif dan negatif) memiliki 12 data.

Tabel 1. Hasil data film threshold 0.3

Nama film	Pistol	Pisau	Rokok	Sentimen
102 Dalmatians	3,164	1,124	816	positif
Babe	1,911	845	669	positif
Christoper R.	1,635	590	967	positif
Dog Purpose	2,093	370	329	positif
Nemo	1,506	192	284	positif
Hachiko	2,136	815	461	positif
Little Stuart 2	1,244	643	704	positif
Marley and Me	2,648	1,008	1,012	positif
Milo and Oits	1,546	212	115	positif
Paddington 2	1,533	1,109	371	positif
Wonder	2,490	857	1,149	positif
Inside Out	2,467	1,187	645	positif
24 Hours Live	3,183	335	500	negatif
Acts Vengeance	2,323	696	428	negatif
Bright	2,576	164	481	negatif
Dead Rising	3,106	365	737	negatif
Final Score	2,819	309	681	negatif
Hitman 47	2,846	634	553	negatif
H. Bodyguard	2,905	266	605	negatif
Looper	2,388	573	756	negatif
Polar	4,184	620	933	negatif
Equalizer 2	2,827	303	445	negatif
The Raid 2	3,971	834	1,141	negatif
Wanted	3,989	608	739	negatif

Pengambilan data dilakukan terhadap semua data film satu persatu menggunakan *darknet.exe* menggunakan 2 *threshold* yang berbeda untuk pengujian akurasi data terbaik yaitu 0,3 yang ditunjukkan Tabel 1 dan 0,5 yang ditunjukkan Tabel 2. Pada kedua tabel tersebut terdapat data pistol, pisau, dan rokok yang menunjukkan jumlah pistol, pisau, dan rokok yang terdeteksi pada film tersebut (deteksi dilakukan per *frame*). Sementara sentimen diisi berdasarkan kode rating umur dari film tersebut. Semua hasil dari pemrosesan data film menggunakan YOLO disimpan dalam bentuk *file csv*.

Saat model YOLO berhasil dilatih melewati 50,000 iterasi dan mendapatkan *loss average* sebesar 0,3615, langkah selanjutnya adalah untuk menguji kekuatan dan akurasi

ISSN 2808-7771 online

dari model YOLO tersebut. Pengujian kekuatan dan akurasi model dilakukan dengan data *test* yang berjumlah 632 dan total objek pada seluruh gambar tersebut ada sebanyak 716 objek. Pengujian untuk mendapatkan akurasi dari model YOLO meliputi *Average Precision* terhadap setiap objek atau kelas, *precision, recall,* dan f1-*score* dari model, dan juga rata-rata *Intersection Over Union* (IOU) yang didapatkan dari model. Semua aspek yang telah disebutkan akan diuji dengan berbagai *threshold* yang berbeda. Ada 5 *threshold* berbeda yang akan digunakan yaitu mulai dari *threshold* 0,3 sampai 0,7.

Tabel 2. Hasil pengujian AP

No.	Object	Average Precision (AP)
1	Pistol	82,80%
2	Pisau	42,11%
3	Rokok	76,64%

Pada hasil pengujian *average precision* (AP) terhadap masing-masing objek yang ditunjukkan pada Tabel 2. Dari hasil tersebut objek pistol mendapatkan AP sebesar 82.80% yang merupakan AP tertinggi dari ketiga objek tersebut. Walaupun data *test* untuk pistol jauh lebih banyak daripada kedua data objek lainnya, AP yang didapatkan masih tinggi dan hal tersebut membuat akurasi model YOLO terhadap deteksi objek pistol dapat dikatan cukup baik. Lalu pada pisau AP yang didapatkan sangat kurang dan paling kecil dari 3 objek yang ada yaitu sebesar 42,11%. Dari hasil AP yang didapatkan oleh objek pisau, maka model YOLO yang telah dilatih masih belum cukup akurat untuk mendeteksi objek pisau dari data *test* yang sudah disiapkan. Rendahnya AP dari objek pisau dapat dikarenakan oleh banyaknya jenis pisau ataupun senjata tajam yang ada dan data yang digunakan belum cukup untuk mewakili seluruh objek pisau atau senjata tajam dalam sebuah film ataupun data *test*. Sedangkan pada objek rokok AP yang didapatkan cukup tinggi yaitu sebesar 76,64%. Berdasarkan AP tersebut, deteksi objek terhadap rokok dalam data *test* bisa dikatakan cukup akurat dan baik. Dari ketiga AP yang didapatkan, maka bisa didapatkan *Mean Average Precision* (mAP) yaitu sebesar 67,18%.

Tabel 3. Hasil pengujian IOU dan confusion matrix

Threshold	TP	FP	FN	Average IOU
0,3	415	65	301	64,65%
0,4	391	54	325	66,01%
0,5	365	44	351	67,26%
0,6	338	34	378	68,81%
0,7	302	23	414	71,24%

Hasil pengujian *threshold* untuk melihat *average* IOU dan *confusion matrix* ditunjukkan pada Tabel 3. Pada tabel tersebut ditunjukkan bahwa *threshold* yang paling kecil yaitu 0,3 mendapatkan cukup banyak *True Positive* (TP) yaitu 415 dari 716 objek, tetapi *threshold* tersebut memiliki total 65 *False Positive* (FP), 301 *False Negative* (FN), dan *average* IOU yang didapatkan *threshold* 0,3 adalah 64,65%. Lalu seiring dengan bertambahnya *threshold* dari 0,4 sampai 0,7 terlihat bahwa TP yang didapatkan menjadi lebih sedikit, FP menjadi lebih sedikit, dan FN menjadi lebih banyak. Hal itu disebabkan karena semakin tinggi *threshold* yang diberikan maka semakin tinggi standar model untuk mendeteksi sebuah objek. Jika *threshold* yang diberikan kecil (0,3) maka model dapat mendeteksi banyak objek karena *threshold* atau probabilitas minimum yang dibutuhkan sangat kecil (0,3). Jika model mendeteksi sebuah objek dan keyakinan atau probabilitas deteksinya diatas 0,3 maka deteksi tersebut adalah valid. Jika *threshold* lebih tinggi (0,7) maka model akan lebih sulit untuk mendeteksi objek karena objek yang dideteksi harus memiliki probabilitas atau keyakinan diatas 0,7 untuk dianggap valid. Tingginya *threshold* sangat berdampak terhadap penurunan TP dan naiknya FN, tetapi FP menjadi menurun.

KESIMPULAN

Model YOLO yang dilatih dapat digunakan untuk mengambil data objek dalam sebuah film dengan dengan akurasi mAP sebesar 67,18%. Akurasi tertinggi didapatkan melalui pendeteksian objek pisto sebesar 82,8% sementara untuk pendeteksian objek pisau sebesar 42,1%. Banyaknya variasi pada gambar pisau di data latih menyebabkan rendahnya akurasi untuk objek pisau.

ISSN 2808-7771 online

DAFTAR PUSTAKA

Aggarwal, CC. (2015). *Data Mining*, Springer, New York. Amerika Serikat.

Du J. (2018). Understanding of Object Detection Based on CNN Family and YOLO, *IOP, 1004 (2018) 012029*.

Komalawati, E. (2017). Industri Film Indonesia: Membangun Keselarasan Ekonomi Media Film dan Kualitas Konten, *ISSN*, 2580-8338.

Larose, T dan Daniel, T. (2005). *Discovering Knowledge in Data : An Introduction to Data Mining*, John Willey & Sons, Hoboken. Amerika Serikat.

Rosebrock, A. (2017). *Deep Learning For Computer Vision With Python, 1th edition*, Pyimagesearch, Hartford. Amerika Serikat.

Redmon, J, Divvala, S, Girshick, R, Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *The IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), 779-788

Redmon, J dan Farhadi, A 2016, *YOLO 9000: Better, Faster, Stronger*, Diakses pada 7 Februari 2019, <https: //pjreddie.com/media/files/papers/YOLO9000.pdf>.

Redmon, J dan Farhadi, A 2018, *YOLOv3: An Incremental Improvement*, Diakses pada 8 Februari 2019, <https: //pjreddie.com/media/files/papers/YOLOv3.pdf>.

© 2021 by authors. Content on this article is licensed under a Creative Commons Attribution 4.0 International license. (http://creativecommons.org/licenses/by/4.0/).