

Sistem Deteksi Senjata Otomatis Menggunakan Deep Learning Berbasis CCTV Cerdas

Iqbal Ahmad Dahlan
Military Engineering Faculty
Indonesia Defense University
Bogor, Indonesia
iqbal.dahlan@idu.ac.id

Dananjaya Ariateja
Military Engineering Faculty
Indonesia Defense University
Bogor, Indonesia
dananjaya.ariateja@idu.ac.id

Muhammad Abditya Arghanie
Military Engineering Faculty
Indonesia Defense University
Bogor, Indonesia
gusti.arghanie@gmail.com

Muhammad Azka Versantariqh
Military Engineering Faculty
Indonesia Defense University
Bogor, Indonesia
azkaesmana8@gmail.com

Muhammad David
Military Engineering Faculty
Indonesia Defense University
Bogor, Indonesia
muhd.davidd@gmail.com

Uvi Desi Fatmawati
Military Engineering Faculty
Indonesia Defense University
Bogor, Indonesia
uvi.fatmawati@idu.ac.id

Abstract— Nowadays, security and safety are big concerns in this modern and cyberwar era. Many countries invest some safety infrastructure to ensure their inhabitants for keeping their lives safely. Indonesia is the country with many problems because of urbanization and other challenges. This problem should be solved with smart city solution and it must be able to face the challenge of ensuring the safety and improving the quality of life regarding network centric warfare era. This problem also should be tackled with CCTV analytics with the ability to implement an automatic weapon detection system. It also can provide the early detection of potentially violent situations that is of paramount importance for citizens security. This paper is using deep Learning techniques based on Convolutional Neural Networks (CNN) can be trained to detect this type of object with YOLOv4 model and it proposes to implement CCTV analytics as a platform to process real-time data for monitoring weapon detection into knowledge displayed in a dashboard with accuracy 0.89, precision 0.82, recall 0.96 dan F1 Score 0.90 result on weapon detection with a real time speed of processing with NVIDIA 2080 Ti around of 35 FPS. It will send an early warning notification if the system detects the weapon detection such as a knife, gun etc.

Keywords—weapon detection, YOLOv4, deep learning, cyberwar, early warning, CNN

I. PENDAHULUAN

Indonesia sebagai salah satu di antara negara berdaulat yang diakui dunia tentunya harus memiliki fungsi dan tujuan negara. Tujuan negara Republik Indonesia seperti yang tercantum dalam teks pembukaan UUD 1945 alinea keempat disebutkan bahwa tujuan negara Indonesia adalah melindungi segenap Bangsa Indonesia dan seluruh tumpah darah Indonesia, memajukan kesejahteraan umum, mencerdaskan kehidupan bangsa dan ikut melaksanakan ketertiban dunia yang berdasarkan kemerdekaan, perdamaian abadi, dan keadilan sosial. Semangat bela negara juga tercantum menurut UU No. 3 Tahun 2002 tentang Pertahanan Negara disebutkan bahwa sikap dan perilaku warga negara yang dijiwai oleh kecintaannya kepada Negara Kesatuan Republik Indonesia yang berdasarkan Pancasila dan Undang-Undang Dasar 1945 dalam menjalin kelangsungan hidup bangsa dan negara yang seutuhnya[1]. Sehingga sebagai warga negara kita juga diwajibkan untuk membela negara sedangkan negara juga berkewajiban untuk memastikan fungsi dan tujuan negara terwujud. Oleh karena itu, sistem pertahanan keamanan rakyat semesta tentunya harus melibatkan seluruh warga

negara dan sumber daya nasional untuk menegakkan kedaulatan negara, keutuhan wilayah, dan keselamatan segenap bangsa dari segala ancaman.

Tantangan era globalisasi juga harus segera diselesaikan dikarenakan permasalahan negara yang semakin kompleks seperti masalah kemiskinan, kesenjangan sosial, banjir, keamanan dan ketertiban perlu dilakukan dengan inovasi dan kebijakan yang tepat dan cerdas. Indonesia sebagai salah satu negara dengan tingkat populasi terbanyak keempat dunia tentunya memiliki masalah yang sangat kompleks selain adanya bonus demografi yang dapat dimanfaatkan namun jika tidak terserap maka akan menjadi sisi negative untuk negara itu sendiri. Selain itu, pertambahan penduduk telah menjadi masalah umum di area perkotaan, baik di dunia maupun di Indonesia. Salah satunya adalah di Jakarta sebagai salah satu kota metropolitan dengan jumlah penduduk terpadat di Indonesia berdasarkan data Dinas Kependudukan dan Pencatatan Sipil, jumlah penduduk DKI Jakarta tahun 2019 sebesar 11.063.324 jiwa yang terbagi dari enam wilayah yaitu Jakarta Pusat, Jakarta Barat, Jakarta Timur, Jakarta Selatan, Jakarta Utara, dan Kepulauan Seribu. Jakarta Pusat sebagai kota pusat pemerintahan menjadi wilayah dengan kepadatan penduduk tertinggi diatas rata-rata sebesar 23.877 jiwa/km²[2]. Bahkan diperkirakan pada tahun 2035 angka tersebut akan meningkat menjadi 66,66% serta meningkat lebih tajam sebanyak 82% pada tahun 2045. Pertumbuhan ini dapat disebabkan oleh tingginya angka kelahiran dan juga disebabkan oleh perpindahan masyarakat dari desa ke kota (urbanisasi) sehingga permasalahan yang ada semakin terpusat di kota. Pertumbuhan populasi di wilayah perkotaan ini akan membuat upaya pengelolaan kota dan menjamin *quality of life* penduduk yang tinggal di kota tersebut menjadi semakin kompleks.

Menurut Data Polda Metro Jaya yang dilansir oleh databoks di tahun 2019 mencatat bahwa DKI Jakarta menjadi daerah yang paling banyak terjadi kasus kejahatan di Indonesia dengan jumlah kejahatan yang dilaporkan ke Metro Jaya mencapai 34,7 ribu kejadian. Jawa Timur menyusul dengan jumlah kejahatan yang dilaporkan mencapai 32,9 ribu kejadian. Permasalahan ini menyebabkan banyaknya keluhan masyarakat di Jakarta, terutama wilayah Jakarta Pusat.

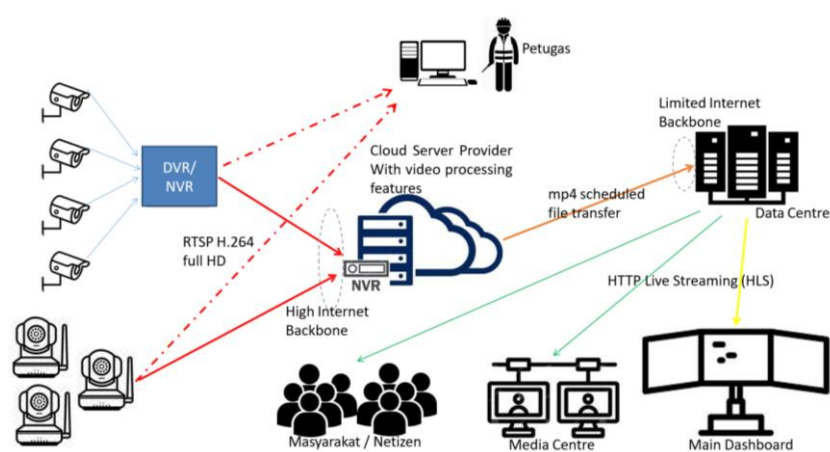
Kondisi tersebut menuntut pemerintah daerah agar memaksimalkan potensi wilayah yang dimiliki dan meminimalisir masalah yang akan dihadapi[3].

Oleh karena itu, solusi konvensional tidak lagi dapat menyelesaikan persoalan kota yang semakin hari semakin kompleks sehingga diperlukan upaya inovasi solusi-solusi cerdas dan inovatif untuk menyelesaikan permasalahan kota dengan efektif dan efisien dengan sistem Smart City yang dibangun tentunya dapat memberikan solusi dan beradaptasi di lingkungan apapun bahkan dalam kondisi di era pandemi seperti saat ini. Berbagai masalah di atas tentunya harus dicari solusi yang tepat dan efektif sehingga masalah tersebut dapat teratasi. Dengan demikian, kami merancang suatu sistem inovasi menggunakan sistem kecerdasan buatan (*artificial intelligence*) berbasis *Video Analytics* dengan memanfaatkan CCTV yang dimiliki pemerintah kota, kepolisian, dan lain-lain untuk memastikan wilayah tersebut aman dan nyaman dengan kemampuan pendeteksi senjata tajam, senjata api dan benda berbahaya lainnya agar secara otomatis dapat terdeteksi. Sistem ini memastikan keamanan dan ketertiban dapat diawasi secara *real time* sehingga kasus di atas bisa diatasi secara cepat ketika sistem mendapatkan suatu anomali atau kejadian yang tidak diinginkan. Contohnya, ada sebuah penelitian yang menggunakan kamera CCTV untuk mengukur angka kriminalitas di kota Gwangmyeong Provinsi Gyeonggi Korea Selatan yang menyimpulkan bahwa pemasangan kamera CCTV di daerah tersebut membuat angka kriminalitas menurun sebesar 47,4%. Bahkan, daerah tetangganya mengalami penurunan angka kriminalitas akibat adanya kamera CCTV di daerah tersebut. Jika dibandingkan dengan daerah yang belum terpasang kamera CCTV, angka kriminalitas di daerah itu tidak mengalami penurunan atau perubahan. Inilah yang kami ingin bangun di wilayah Jakarta untuk mengurangi permasalahan yang ada tersebut dengan bantuan teknologi. Sehingga penerapan kamera CCTV ini mampu mengurangi tingkat kriminalitas dan meningkatkan ketertiban di setiap daerah. Apalagi penerapan tersebut menggunakan sistem kecerdasan buatan yang berbasis *Video Analytics* dengan kemampuan deteksi senjata ini akan bisa berperan lebih aktif dalam melakukan pengawasan secara mandiri yang di mana data dari kamera CCTV tersebut akan dihubungkan

langsung ke pusat server/pusat kebijakan (*command center*) dengan cepat dan akurat.

II.Kajian Pustaka

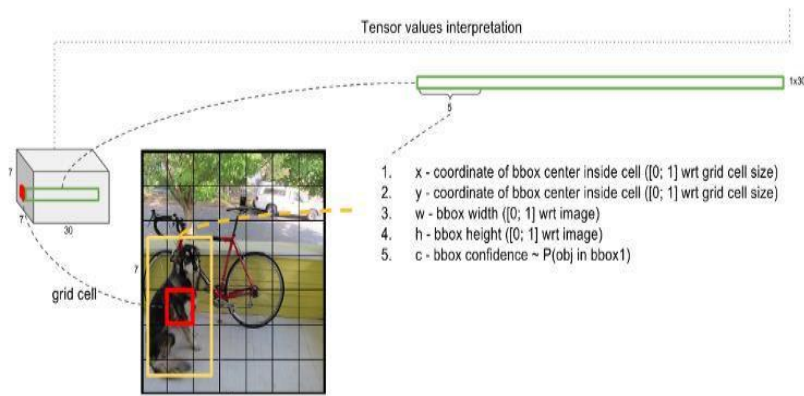
Salah satu sensor yang kita gunakan adalah CCTV sebagai alat yang digunakan untuk memantau apa yang tertangkap oleh alat tersebut yang dikenali dalam format video. Sistem yang dirancang tidak terikat dengan merk CCTV tertentu sehingga dapat menggunakan CCTV yang sudah ada atau yang dimiliki oleh pengguna. CCTV akan diintegrasikan melalui protokol video streaming seperti RTSP, RTMP atau HTTP sehingga dari CCTV dapat dilihat dan dianalisis di sistem lainnya. Tahapan selanjutnya adalah analisis citra yang diambil langsung video melalui protokol yang kami gunakan adalah RTSP dari kamera CCTV berupa gambar per *frame* yang sangat cepat tergantung berapa FPS video yang kita dapatkan[2].



Gambar 1. Skema akuisisi data CCTV[3].

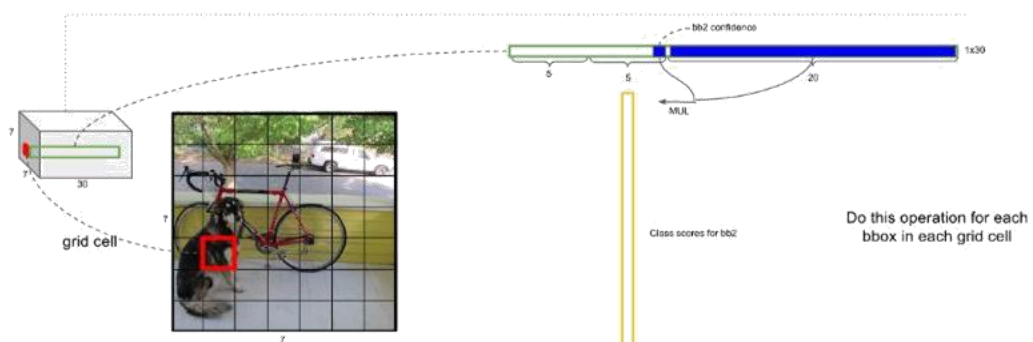
II.1 Deteksi Objek

Sistem pendeteksi modul multi objek didukung oleh algoritma Convolutional Neural Network untuk pengolahan data dan pengenalannya.



Gambar 2. Pemetaan deteksi koordinat objek[4].

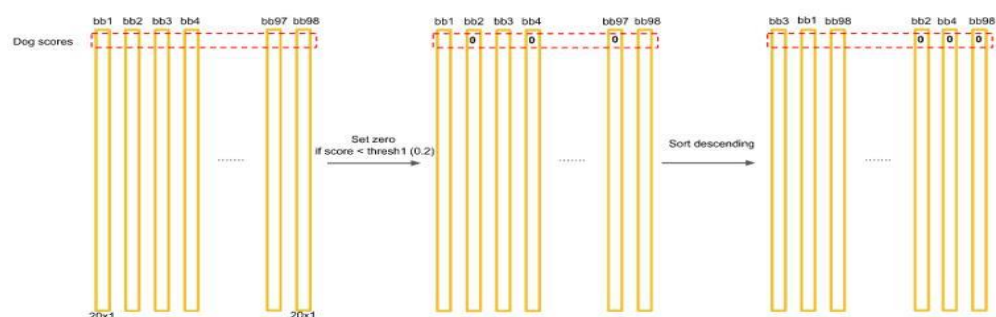
Sistem akan membuat grid cell pada frame dimana objek akan dikenali dan ditandai dengan label box yang mana data itu akan diinterpretasikan kedalam posisi x-y dan membentuk box disekitarnya yang didalamnya terdapat box center objek yang akan kita kenali. Hal ini digunakan untuk mendeteksi jika objek tersebut berdekatan untuk itulah titik tengah per objek menjadi perlu untuk kita deteksi agar setiap objek dapat dikenali. Untuk setiap sel grid, sistem akan mendapatkan dua kotak pembatas, yang akan menjadi nilai awal 10 dari tensor $1 * 30 * 20$ sisanya menunjukkan jumlah kelas. Nilai-nilai menunjukkan nilai kelas, yang merupakan probabilitas bersyarat dari objek milik kelas i , jika suatu objek hadir di dalam kotak. Selanjutnya, kita mengalikan semua nilai kelas dengan nilai bobot kotak dan mendapatkan nilai kelas untuk batas kotak selanjutnya yang berbeda. Step selanjutnya adalah melakukan ini untuk semua sel grid. Itu sama



dengan $7 * 7 * 2 = 98$ sehingga sistem memiliki skor kelas untuk setiap kotak pembatas (dimensi tensor = $20 * 1$).

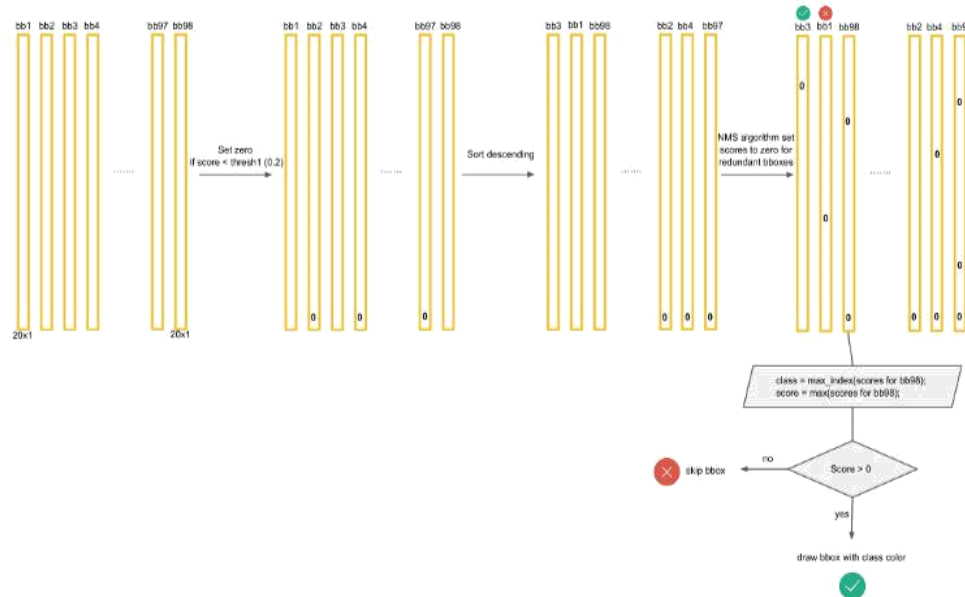
Gambar 3. Pengenalan objek [4].

Sekarang mari kita fokus pada contoh gambar anjing dalam gambar. Skor anjing untuk kotak pembatas akan hadir dalam (1,1) dari tensor di semua skor kotak yang kita amati secara sekuensial akan menetapkan nilai ambang nilai dan mengurutkan mereka secara berurutan.



Gambar 4. Pembobotan box yang diisi objek [4].

Sekarang kita akan menggunakan algoritma supresi Non-maks untuk menetapkan skor ke nol untuk kotak redundan. Pertimbangan Anda memiliki skor anjing untuk boundingbox1 sebagai 0,5 dan biarkan ini menjadi skor tertinggi dan untuk box47 sebagai 0,3. Lalu akan mengambil titik-temu atas persatuan nilai-nilai ini dan jika nilainya lebih besar dari 0,5, akan menetapkan nilai untuk kotak2 sebagai nol, jika tidak, kami akan melanjutkan ke kotak berikutnya lalu system akan melakukan ini untuk semua kotak. Setelah semua ini dilakukan, kita akan menyisakan dengan 2-3 kotak saja. Sedangkan nilai kotak lain akan nol. Sekarang, kami memilih bbox untuk menggambar berdasarkan nilai



nilai kelas. Ini dijelaskan dalam gambar. Inilah yang akan dikenali dengan objek per klaster yang sudah kita bagi sehingga pendeteksian multiobjek menjadi mungkin untuk dideteksi.

Gambar 5. Pengambilan keputusan bobot box [4].

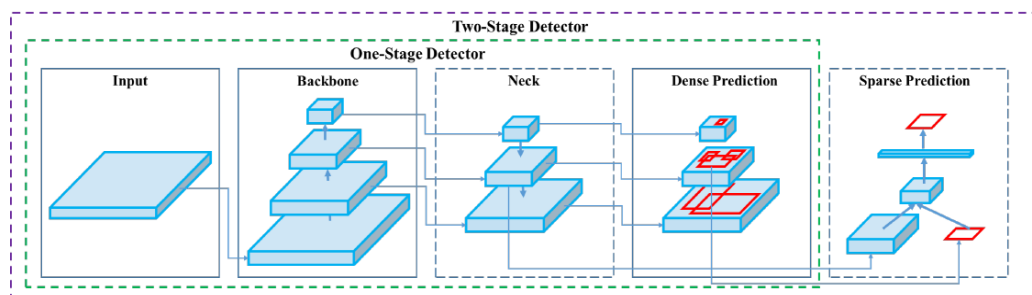
Sistem ini menerapkan model ini menggunakan Convolutional Neural Network /Deep Learning untuk mengevaluasi set data deteksi PASCAL VOC. Layer konvolusional awal dari ekstrak jaringan fitur dari gambar sementara yang terhubung sepenuhnya memprediksi probabilitas output dan koordinat. Arsitektur jaringan kami terinspirasi oleh GoogLeNet model untuk klasifikasi gambar.

Jaringan CNN ini memiliki 24 layer konvolusional CNN diikuti oleh 2 lapisan yang terhubung sepenuhnya.

II.2 Scaled YOLOv4

Algoritma Scaled YOLOv4 adalah salah satu algoritma deteksi objek yang dapat mengetahui objek apa yang dideteksi oleh kamera. Secara umum model ini dibagi menjadi 2 bagian yaitu *backbone* dan bagian *head*. *Backbone* digunakan secara umum untuk melatih data pada ImageNet dan *head* memiliki fungsi untuk memprediksi *class* dan *bounding box* pada objek tersebut. Pada sisi *head* juga dibedakan menjadi 2 jenis. Jenis pertama detektor objek satu tahap dan detektor objek dua tahap. Detektor objek satu tahap contohnya adalah YOLO, RPN dan RetinaNet dan detektor objek dua tahap contohnya adalah R-CNN dan R-FCN. Detektor objek dikembangkan pada beberapa tahun ini biasanya menambah satu layer diantara *backbone* dan *head* dengan tujuan pengumpulan beberapa *feature map* di beberapa tahap dikenal dengan kata *neck layer*. Secara umum, neck layer terdiri dari beberapa bagian seperti FPN, PAN, BiFPN, dan NAS-FPN. Berikut ini dijelaskan mengenai arsitektur pada detektor objek yang dikembangkan [5][6].

- **Input** : Image, Patches, Image Pyramid.
- **Backbones** : EfficientNet-B0/B7, CSPResNeXt50, CSPDarknet53 VGG16, ResNet-50, SpineNet,.
- **Neck** :
 - *Additional Blocks*: SPP, SAM, ASPP, RFB.
 - *Path-aggregation blocks*: FPN, BiFPN, ASFF, SFAM PAN, NAS-FPN, Fully-connected FPN,.
- **Heads** :
 - *Dense Prediction (one-stage)* : RPN, CornerNet, CenterNet, MatrixNet, FCOS, SSD, YOLO, RetinaNet.
 - *Sparse Prediction (two-stage)* : Faster R-CNN, R-FCN, Mask R-CNN, RepPoints, R-CNN.



Gambar 7. Perbandingan algoritma / model[6] .

Arsitektur YOLO yang didesain dengan menggunakan Scaled YOLOv4 dapat diatasi dengan focal loss untuk mengatasi permasalahan yang terdapat pada ketidakseimbangan data. Tahapan terakhir adalah IoU loss, metode ini digunakan untuk mengatasi masalah yang terdapat pada fungsi regresi bounding box dalam memastikan integritas objek yang terdeteksi. Implementasi pendeteksian objek pada sistem membutuhkan sumber daya tertentu untuk dapat menghasilkan pendeteksian objek yang diinginkan, untuk penambahan modul dan metode post-processing yang menambah kebutuhan sumber daya yang lebih sedikit pada saat pendeteksian tetapi dapat menghasilkan akurasi pendeteksian yang lebih signifikan disebut dengan *bag of specials*[5][6]. Secara umum modul ini berfungsi untuk memperbaiki beberapa atribut pada model yang dibuat, seperti meningkatkan bidang reseptif, menerapkan mekanisme perhatian, atau meningkatkan kemampuan integrasi fitur, dan metode post-processing berfungsi untuk menyaring hasil prediksi yang telah dibuat oleh modul model. Beberapa modul yang sering digunakan di kantong spesial adalah SPP, ASPP, dan RFB, salah satu hasil yang didapat dari modul ini adalah ketika Redmon menerapkan modul SPP ke YOLOv3, hasil AP yang diperoleh meningkat sebesar 2,7% pada Deteksi objek MS COCO hanya dengan menambahkan perhitungan ekstra sebesar 0,5% pada sistem[6]. Modul Attention yang sering digunakan dalam pendeteksian objek adalah Squeeze-and-Excitation dan Spatial Attention Module, kedua modul ini dapat meningkatkan akurasi model yang dihasilkan dengan mengorbankan beberapa komputasi yang diperlukan. Pada tahun 2010 Nair dan Hinton mengusulkan ReLU untuk memecahkan masalah gradien menghilang yang biasanya ditemukan pada fungsi aktivasi tanh dan sigmoid, sejak itu beberapa iterasi dan perbaikan pada metode ReLU telah dikembangkan untuk memungkinkan model deteksi objek beroperasi lebih efisien. Dalam post-processing, metode yang sering digunakan saat ini untuk deep learning adalah Non Max Suppression, yaitu metode yang digunakan untuk menyaring BBox yang tidak dapat memprediksi objek yang sama dengan baik, sehingga hanya kandidat BBox yang akurat yang akan ditampilkan pada hasil deteksi akhir. Metode ini juga terus dikembangkan dari waktu ke waktu agar hasil

pendeteksian objek dapat lebih baik dan implementasi sistem dapat dilakukan secara efisien. Secara umum, Scaled-YOLOv4 memegang beberapa konsep penskalaan secara seimbang saat peneliti mengerjakan konstruksi model - ukuran gambar, jumlah lapisan, dan jumlah saluran, sambil mengoptimalkan kinerja model dan kecepatan inferensi. Para peneliti mempertimbangkan beberapa tulang punggung CNN berukuran CSP dalam jaringan, ResNet, ResNeXt, dan Darknet backbone sebelumnya[6].

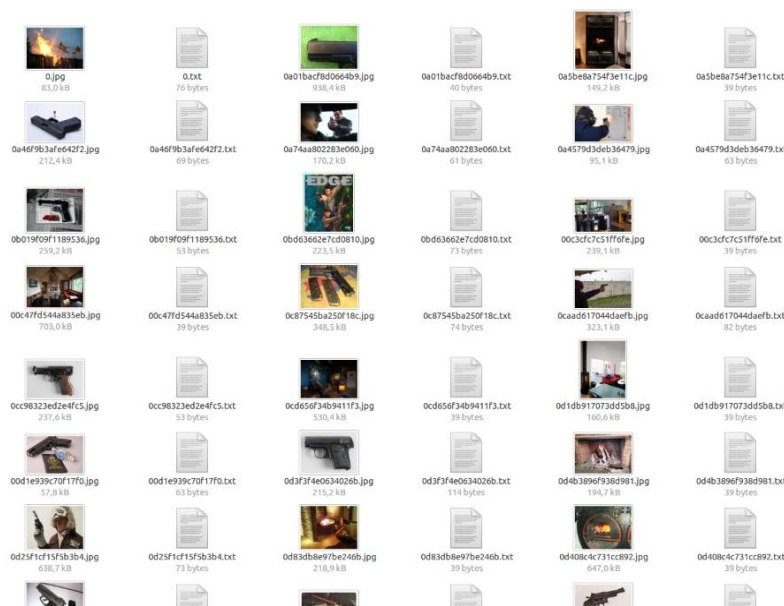
Dalam makalah Scaled-YOLOv4 peneliti sering menulis bahwa mereka melakukan Cross-Stage-Partial (CSP) bagian tertentu dari jaringan. Mengukur CSP berarti menerapkan konsep yang tercantum dalam makalah Jaringan Parsial Lintas Tahap, yang ditulis oleh WongKinYiu. CSP adalah cara baru untuk merancang jaringan saraf convolutional yang menghemat komputasi untuk berbagai jaringan CNN - hingga 50% (untuk backbone Darknet di FLOP)[5][6].

Untuk mendeteksi objek besar dalam gambar besar, peneliti menemukan bahwa penting untuk meningkatkan kedalaman dan jumlah tahapan di dalam *backbone* dan *neck layer* CNN. (Meningkatkan lebar dilaporkan memiliki sedikit efek.) Hal ini memungkinkan untuk terlebih dahulu meningkatkan ukuran input dan jumlah tahapan, dan secara dinamis menyesuaikan lebar dan kedalaman sesuai dengan persyaratan kecepatan inferensi waktu nyata. Selain faktor penskalaan ini, penulis juga menyesuaikan konfigurasi arsitektur model mereka di makalah. Evaluasi menunjukkan bahwa model Scaled-YOLOv4 menetapkan standar untuk model deteksi objek besar dan kecil[5].

sistem ini perlu diperbaiki dalam integrasi ke dashboard agar dapat memudahkan pengguna dalam melihat kejadian orang membawa senjata berbahaya tersebut.

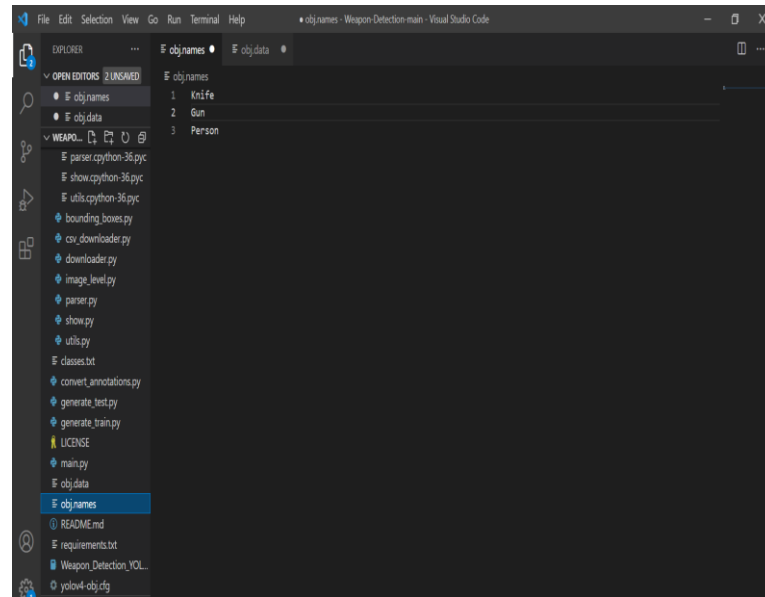
III Implementasi Riset

Hal pertama yang dilakukan adalah pengumpulan data senjata seperti pisau, pistol dan senjata lainnya. Hal inilah yang nantinya kita latih sehingga nantinya akan ditemukan model berdasarkan data yang sudah kita latih. Data dapat diambil secara bebas di laman (<https://www.kaggle.com/atulyakumar98/fire-and-gun-dataset>) [7] dan <https://dasici.es/transferencia/open-data/24705/>.



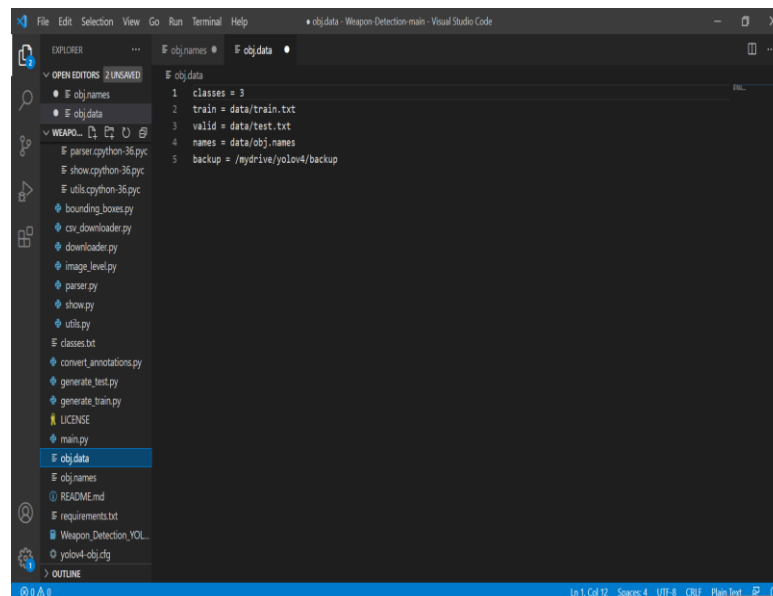
Gambar 7. Dataset senjata [7].

Buat file obj.data dan obj.names untuk melatih dataset yang sudah kita miliki. File obj.names akan terdiri dari 3 kategori objek yaitu orang dan senjata pisau (*knife*) dan pistol (*gun*) .



Gambar 8. Konfigurasi obj.names.

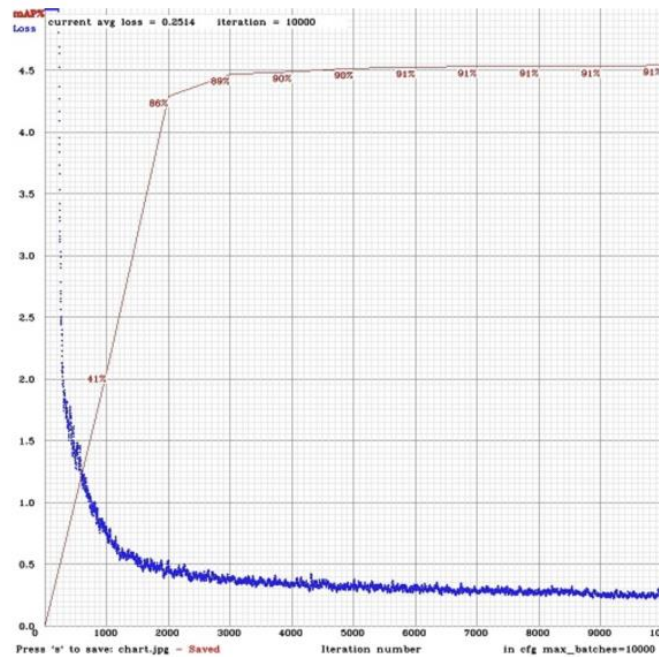
Konfigurasi obj.data adalah sebagai berikut sesuai dengan Gambar 9. Proses inilah yang nantinya akan melatih dataset yang kita punyai sehingga pada akhirnya kita peroleh model dari data yang sudah kita miliki sesuai dengan jumlah kelas yang kita sudah atur dalam kasus ini 3 kelas yaitu orang (*Person*) atau senjata pisau dan pistol.



Gambar 9. Konfigurasi obj.data.

Data tersebut akan dilatih menggunakan pre-trained model yang sudah kita miliki dengan mendownload dengan halaman berikut https://github.com/AlexeyAB/darknet/releases/download/darknet_yolo_v3_optimal/yolov4.conv.137. Proses pelatihan akan dilakukan dan akan disimpan di folder

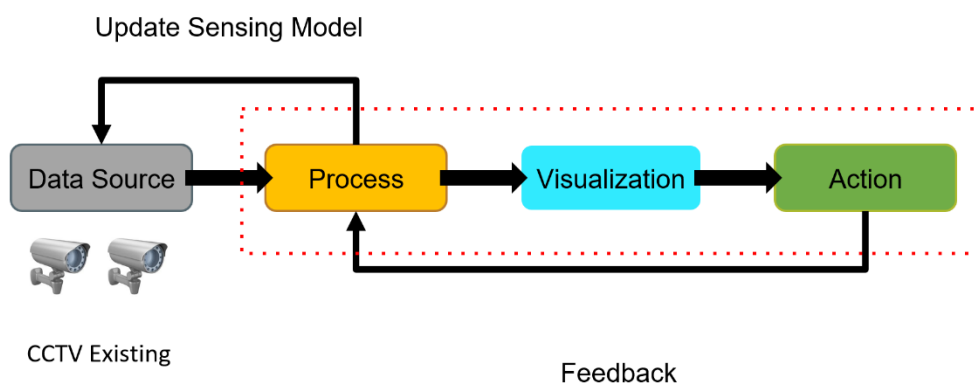
/mydrive/yolov4/backup untuk menemukan model akhir. Proses pelatihan ini akan ditunjukkan ke dalam gambar berikut dan didapatkan avg loss sebesar 0.2514:



Gambar 10. Grafik latihan model.



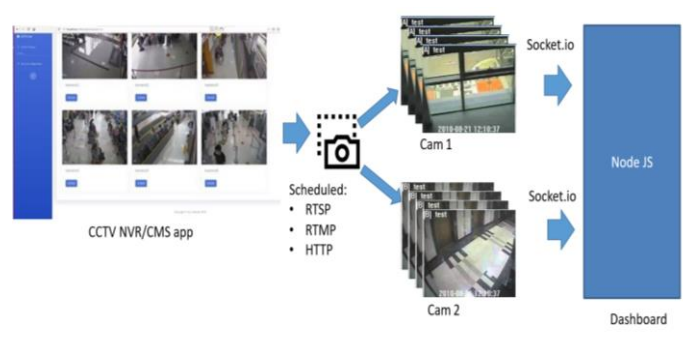
Gambar 11. Prediksi deteksi senjata



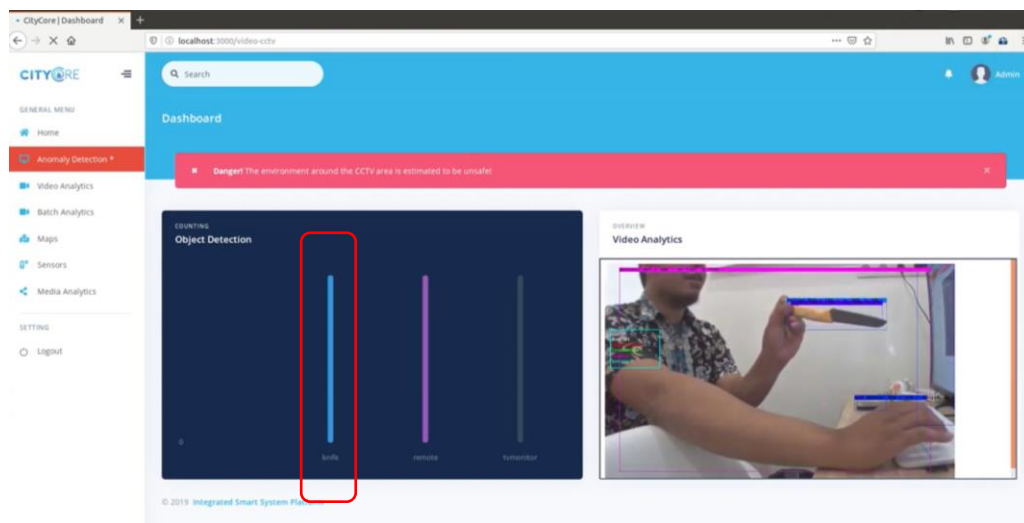
Gambar 12. Arsitektur Sistem CCTV.

IV Finalisasi Sistem

Sistem membutuhkan dashboard untuk membuat stakeholder dalam implementasi sistem deteksi senjata ini. Sistem akan memonitor secara real time analisis dan akan memberikan notifikasi jika sistem cctv menemukan pengunjung / orang yang membawa senjata tajam. Sistem ini menggunakan Nodejs untuk *frontend dashboard* dan MongoDB sebagai database NoSQL nya[6] [7].



Gambar 13. Skema akuisisi data CCTV ke dashboard.



Gambar 14. Weapon Detection Dashboard.

Sistem dicoba dengan IMFDB dataset diperoleh hasil berikut accuracy 0.83, precision 0.80, recall 0.88 dan F1 Score 0.83 dan juga dibandingkan dengan UGR Dataset untuk data <https://dasici.es/transferencia/open-data/24705/> diperoleh accuracy 0.89, precision 0.82, recall 0.96 dan F1 Score 0.90.

V Kesimpulan

Paper ini menggunakan algoritma CNN dengan model scaled Yolov4 dan

konsep integrasinya menggunakan teknologi Nodejs sebagai frontend dashboard dan MongoDB sebagai database NoSQL untuk mengatasi permasalahan monitoring kejadian orang membawa senjata tajam atau senjata api yang berbahaya dengan dashboard yang interaktif. Sistem ini bertujuan untuk mendeteksi senjata yang berbahaya otomatis dengan avg loss sebesar 0.2514 dan rata-rata 35 FPS dengan NVIDIA RTX 2080 Ti dengan akurasi dengan IMFDB dataset diperoleh hasil berikut accuracy 0.83, precision 0.80, recall 0.88 dan F1 Score 0.83 dan juga dibandingkan dengan UGR Dataset untuk data <https://dasci.es/transferencia/open-data/24705/> diperoleh accuracy 0.89, precision 0.82, recall 0.96 dan F1 Score 0.90.

Daftar Pustaka

- [1] R. Indonesia, "UU Nomor 3 Tahun 2003," no. September, pp. 1–2, 2009, [Online]. Available: http://www2.pom.go.id/public/hukum_perundangan/pdf/.
- [2] I. A. Dahlan, F. Hamami, S. H. Supangkat, and F. Hidayat, "Big Data Implementation of Smart Rapid Transit using CCTV Surveillance," *Proceeding - 2019 Int. Conf. ICT Smart Soc. Innov. Transform. Towar. Smart Reg. ICISS 2019*, pp. 1–5, 2019, doi: 10.1109/ICISS48059.2019.8969830.
- [3] B. D. Riyanto, "Cloud Service Design for Computer Vision , Image / Video Processing & Analytics," no. June, 2021.
- [4] J. Redmon and A. Farhadi, "YOLO v.3," *Tech Rep.*, pp. 1–6, 2018, [Online]. Available: <https://pjreddie.com/media/files/papers/YOLOv3.pdf>.
- [5] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "Scaled-YOLOv4: Scaling Cross Stage Partial Network," 2020, [Online]. Available: <http://arxiv.org/abs/2011.08036>.
- [6] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv*, 2020.
- [7] P. Mehta, A. Kumar, and S. Bhattacharjee, "Fire and Gun Violence based Anomaly Detection System Using Deep Neural Networks," *Proc. Int. Conf. Electron. Sustain. Commun. Syst. ICESC 2020*, no. Icesc, pp. 199–204, 2020, doi: 10.1109/ICESC48915.2020.9155625.

- [8] M. T. Bhatti, M. G. Khan, M. Aslam, and M. J. Fiaz, "Weapon Detection in Real-Time CCTV Videos Using Deep Learning," *IEEE Access*, vol. 9, pp. 34366–34382, 2021, doi: 10.1109/ACCESS.2021.3059170.
- [9] F. Hidayat, F. Hamami, I. A. Dahlan, S. H. Supangkat, A. Fadillah, and A. Hidayatuloh, "Real Time Video Analytics Based on Deep Learning and Big Data for Smart Station," *J. Phys. Conf. Ser.*, vol. 1577, no. 1, 2020, doi: 10.1088/1742-6596/1577/1/012019.