

BAB 1

PENDAHULUAN

1.1. Latar Belakang

Isu krusial yang sering menjadi fokus perhatian di berbagai negara salah satunya adalah keamanan publik. Tindakan kriminal, salah satunya adalah kekerasan, dapat terjadi kapan saja dan di mana saja, baik di ruang publik maupun pribadi seperti tawuran, pembegalan, *bullying*, dan aniaya. Kekerasan dapat mempengaruhi kehidupan individu, keluarga, dan komunitas di seluruh dunia. Dampak negatifnya meliputi kecanduan zat, depresi, bunuh diri, putus sekolah, pengangguran, dan kesulitan interpersonal yang berulang (Jacob K.S, 2016). Menurut data dari Kementerian Pemberdayaan Perempuan dan Perlindungan Anak (Kemen-PPPA), Antara 1 Januari hingga 27 September 2023, tercatat sebanyak 19.593 kasus kekerasan di Indonesia. Data ini bersifat *real-time* dan diperbarui hingga pukul 14.35 WIB pada saat pelaporan. Dari data kasus kekerasan yang terjadi berdasarkan periode tersebut, 3.987 korban berjenis kelamin laki-laki dan 17.347 adalah perempuan. Berdasarkan usia, sebanyak 38% dari data periode tersebut didominasi oleh kelompok pada rentang usia 13-17 tahun, yang mencapai 7.451 korban. (Nabilah Muhamad, 2023)

Upaya penanganan kekerasan telah diterapkan oleh pemerintah kerap kali belum terlaksana dengan baik seperti kebijakan dan program yang bertujuan untuk melindungi korban serta mencegah terjadinya kekerasan lebih lanjut. Pemerintah dan lembaga-lembaga terkait sering kali menyediakan layanan dukungan psikologis, perlindungan hukum, serta akses ke tempat penampungan untuk para korban. Namun, di tengah meningkatnya jumlah kasus kekerasan, tantangan yang muncul adalah bagaimana memberikan respon cepat dan tepat sasaran terhadap kejadian-kejadian tersebut. Dalam konteks ini, memanfaatkan kamera pengawas seperti CCTV dan teknologi model *deep learning* berpotensi memberikan solusi secara *real-time* untuk menjawab permasalahan respon penanganan terhadap tindak kekerasan fisik. Pemanfaatan model *deep learning* dalam kamera pengawas dapat memberikan prediksi adanya kekerasan dan memungkinkan tindakan yang lebih cepat dan tepat dari pihak keamanan, sehingga meningkatkan efektivitas dalam merespon insiden kekerasan.

Penelitian sebelumnya oleh Sharma *et al* yang berjudul *A fully integrated violence detection system using CNN and LSTM* pada tahun 2021 memanfaatkan *Xception* sebagai *feature extractor* dan *LSTM* untuk menangkap dan mempelajari fitur sekuensial. Skema train-test yang digunakan yaitu 50 *epochs* dengan 80% data *train* dan 20% data *validation* dengan hasil akurasi mencapai 98.87%.

Model *YOLO* merupakan metode yang sering digunakan untuk deteksi objek. Model ini dikenal karena kemampuannya mendeteksi objek dengan cepat dan efisien dalam satu kali pemrosesan, menjadikannya sangat ideal untuk aplikasi *real-time* (Redmon *et al.*, 2016). *YOLO* bekerja dengan menggunakan pendekatan *single-shot detection* di mana seluruh gambar hanya diolah dalam satu langkah pemrosesan lalu membagi gambar *input* menjadi beberapa *grid* kecil dan setiap sel *grid* bekerja untuk memprediksi *bounding boxes* dan *confidence score* yang menunjukkan seberapa besar kemungkinan objek tertentu ada di dalam *bounding boxes* tersebut. Penelitian sebelumnya efektivitas *YOLO* menunjukkan tingkat akurasi mencapai *mAP* (*mean Average Precision*) sebesar 57.9% pada *dataset COCO*, dan model ini mampu memproses gambar pada kecepatan hingga 45 *frame* per detik pada versi *YOLOv3* dalam berbagai aplikasi, termasuk deteksi objek dan klasifikasi aktivitas (Redmon & Farhadi., 2018).

MobileNetv3 bekerja menggunakan teknik *depthwise separable convolutions*, di mana setiap *channel input* diproses secara terpisah menggunakan *kernel* konvolusi. Teknik ini membuat *MobileNetv3* lebih ringan karena memanfaatkan sistem paralelisme, di mana setiap *channel* diproses secara terpisah namun bersamaan, sehingga mengurangi jumlah operasi dibandingkan dengan konvolusi tradisional. *MobileNetv3* juga memperkenalkan beberapa peningkatan arsitektur, seperti penggunaan non-linearitas baru yang lebih efisien (*h-swish*) dan desain ulang lapisan awal dan akhir jaringan.

1.2. Perumusan Masalah

1. Bagaimana penggabungan model *Yolov11* untuk mendeteksi manusia dengan model *MobileNetv3* untuk memprediksi adanya kekerasan dalam video
2. Bagaimana mengatasi masalah ketidakseimbangan data (*imbalanced data*) antara *frame* yang mengandung kekerasan dan *frame normal*, sehingga model tidak bias terhadap kelas mayoritas

1.3. Tujuan Penelitian

Penelitian ini memiliki tujuan utama yaitu mengembangkan model prediksi kekerasan berbasis video dengan menggabungkan model *Yolov11* untuk deteksi manusia dan *MobileNetv3* untuk klasifikasi kekerasan.

1.4. Batasan Masalah

Agar penelitian ini lebih terfokus dan terarah, beberapa batasan masalah telah ditetapkan. Pembatasan ini dilakukan berdasarkan ruang lingkup penelitian serta kemampuan teknis dalam aspek yang berkaitan dengan pengembangan sistem deteksi kekerasan berbasis video. Batasan-batasan yang ditetapkan adalah sebagai berikut:

1. Jenis Deteksi dan Klasifikasi

Penelitian ini hanya fokus pada dua jenis deteksi: deteksi keberadaan manusia menggunakan model *Yolov11* dan mendeteksi adanya aktivitas kekerasan yang bersifat prediktif. Kekerasan yang dideteksi pada penelitian hanya prediksi sebatas kekerasan fisik tanpa prediksi senjata tajam/tumpul/api dan tidak sampai pada prediksi kekerasan psikis. Deteksi objek selain manusia tidak akan dibahas pada penelitian ini.

2. Dataset

Dataset yang digunakan terdiri dari video kekerasan dan non-kekerasan yang telah diklasifikasikan sebelumnya oleh pakar Widyaiswara. Dataset terdiri dari dataset primer dan sekunder. Data sekunder diambil dari sumber terbuka dan dataset primer dibuat oleh peneliti langsung dengan skenario gerakan kekerasan tertentu dan keduanya telah melewati proses validasi oleh pakar konseling. Dataset primer hanya diambil pada kamera pengawas yang bisa diakses peneliti.

3. Lingkup Implementasi

Penelitian ini membatasi implementasi sistem deteksi kekerasan berbasis video hanya pada situasi simulasi yang dijalankan dalam lingkungan komputer pribadi (*PC/laptop*) dengan menggunakan video input standar. Sistem ini tidak akan diuji dalam skala besar seperti penerapan model pada perangkat pengawasan video dari *CCTV* di tempat umum.

4. Kualitas Video

Kualitas video yang digunakan dalam penelitian ini dibatasi pada resolusi standar maksimal (1080p) atau lebih rendah.

5. Lingkungan Pengujian

Data primer dibuat pada lingkungan dengan pencahayaan stabil dan dua sudut pandang.

6. Waktu Pengolahan

Bentuk prediksi adalah *Sequence-to-one* dimana hasil prediksi adalah kesimpulan dari fitur sekuensial yang diekstrak (prediksi video), sehingga prediksi *per-frame* tidak dilakukan. Optimasi perangkat keras untuk menjalankan model tidak akan dibahas pada penelitian ini.

7. Objek Aktivitas

Penelitian ini berfokus pada aktivitas yang secara umum diidentifikasi sebagai kekerasan fisik seperti pukulan, tendangan, dan serangan fisik lainnya. Gerakan yang menyerupai kekerasan namun memiliki konteks budaya tertentu, seperti salam berpelukan atau tarian modern/tradisional, dan ekspresi wajah tidak menjadi bagian dari ruang lingkup penelitian ini.

8. *Sequence Length*

Dalam penelitian ini, *Sequence Length* dibatasi hingga 16. Hal ini didasarkan pada hasil eksperimen dengan nilai 32, 64, dan 128, yang menunjukkan bahwa semakin besar *sequence length*, semakin tinggi kebutuhan memori dan waktu pelatihan. Pada percobaan dengan nilai lebih besar, model mengalami *crash* akibat keterbatasan memori, terutama saat dijalankan pada *cloud notebook Kaggle*. Oleh karena itu, *Sequence Length* = 16 dipilih sebagai nilai optimal untuk menjaga keseimbangan antara kinerja model dan efisiensi komputasi.

Batasan masalah ini dilakukan untuk memastikan bahwa penelitian tetap terfokus pada pengembangan dan pengujian sistem deteksi kekerasan secara efisien dengan teknologi yang tersedia, dan agar hasil penelitian dapat diselesaikan sesuai dengan waktu dan sumber daya yang dimiliki.

1.5. State of the Art

Judul Jurnal	Pembahasan
<p><i>Lightweight Network for Violence Detection Using YOLOv5 and MobileNet</i></p> <p>Peneliti: YoushanZhang¹, YongLi², ShaozheGuo¹</p> <p>Lokasi:</p>	<p>Hasil Penelitian:</p> <p>Mengembangkan model deteksi kekerasan yang ringan dan efisien dengan menggabungkan <i>YOLOv5</i> dan <i>MobileNet</i>, memungkinkan aplikasi di perangkat <i>mobile</i> dan <i>embedded</i>.</p>

<p>1. GraduateStudentBrigade, Chinese People's Armed Police ForceEngineering University, Xi'an, Shaanxi Province, China,</p> <p>2. College of Information Engineering, Chinese People's Armed Police Force Engineering University, Xi'an, Shaanxi Province, China</p> <p>Tahun 2022</p> <p>Nama Jurnal PLOS One</p>	<p>Alasan Menjadi Tinjauan Penelitian:</p> <p>Penelitian ini cukup relevan dengan penelitian yang penulis sedang kerjakan terutama pada optimasi penggunaan <i>TSM</i> sebagai bahan inspirasi untuk <i>tuning MobileNet</i>.</p> <p>Penggunaan kombinasi dari <i>YOLOv5</i> dan <i>MobileNet</i>, serta penambahan modul <i>TSM</i>, menunjukkan inovasi dalam arsitektur jaringan yang bertujuan untuk memaksimalkan kinerja sambil meminimalkan biaya komputasi.</p>
<p><i>A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv11 and YOLO-NAS</i></p> <p>Peneliti: Juan Terven, Diana-Margarita Córdova-Esparza, dan Julio-Alejandro Romero-González</p> <p>Lokasi: 1. Instituto Politecnico Nacional, CICATA-Qro, Queretaro76090, Mexico 2. Facultad de Informática, Universidad Autónoma de Querétaro, Queretaro76230, Mexico;</p> <p>Tahun 2023</p> <p>Nama Jurnal</p>	<p>Hasil Penelitian:</p> <p>Analisis mendalam tentang inovasi <i>YOLO</i> dalam deteksi objek, peningkatan akurasi, dan efisiensi komputasi</p> <p>Alasan Menjadi Tinjauan Penelitian:</p> <p>Memberikan wawasan komprehensif tentang perkembangan <i>YOLO</i>, penting sebagai referensi utama dalam penelitian yang menggunakan <i>YOLO</i> untuk deteksi kekerasan atau objek.</p>

MDPI	
<p><i>Research on Pedestrian Detection Algorithm Based on MobileNet-YoLo</i></p> <p>Peneliti: Lisang Liu¹, Chengyang Ke¹, HeLin² dan Hui Xu¹</p> <p>Lokasi: 1. School of Electronic Electrical Engineering and Physics, Fujian University of Technology, Fuzhou, Fujian, China 2. State Grid Fujian Power Supply Co. LTD., Xiapu Power Supply Company, Ningde, Fujian, China</p> <p>Tahun 2022</p> <p>Nama Jurnal Hindawi</p>	<p>Hasil Penelitian: Memperkenalkan sebuah model MobileNet-YOLO yang dirancang untuk mendeteksi pejalan kaki pada perangkat dengan daya komputasi rendah. Model ini dikembangkan dari YOLOv4-tiny, dengan peningkatan pada arsitektur jaringan menggunakan MobileNet untuk ekstraksi fitur yang lebih efisien.</p> <p>Alasan Menjadi Tinjauan Penelitian: Penelitian ini menawarkan pendekatan efisien untuk deteksi objek pada perangkat kecil dengan pengurangan parameter yang signifikan. Metode ini dapat digunakan untuk membangun sistem deteksi kekerasan berbasis MobileNet yang ringan dan cepat.</p>
<p><i>Violence-YOLO: Enhanced GELAN Algorithm for Violence Detection</i></p> <p>Peneliti Wenbin Xu¹, Dingju Zhu^{1,2}, Renfeng Deng¹, KaiLeung Yung³ and Andrew W. H. Ip⁴</p> <p>Lokasi 1. School of Software, South China Normal University, Foshan 528000, China School of</p>	<p>Hasil Penelitian: Memperkenalkan Violence-YOLO, sebuah model yang didesain khusus untuk mendeteksi kekerasan dalam video secara akurat dan <i>real-time</i>, terutama dalam lingkungan kompleks seperti bandara dan ruang publik.</p> <p>Alasan Menjadi Tinjauan Penelitian: Model ini menggabungkan teknik YOLO terbaru dengan peningkatan arsitektur untuk</p>

<p>2. <i>Computer Science, South China Normal University, Guangzhou 510000, China</i></p> <p>3. <i>Department of Industrial and Systems Engineering, Hong Kong Polytechnic University, HongKong999077, China</i></p> <p>Department of</p> <p>4. <i>Mechanical Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A2, Canada</i></p> <p>Tahun 2024</p> <p>Nama Jurnal <i>Applied Science</i>, MDPI</p>	<p>meningkatkan akurasi deteksi kekerasan di lingkungan publik dengan latensi rendah dan performa tinggi. Relevan karena dapat diadaptasi untuk mendeteksi kekerasan dalam <i>video surveillance</i>.</p>
<p><i>Improving Reliability of Fine-tuning with Block-wise Optimisation</i></p> <p>Peneliti: Basel Barakat dan Qiang Huang</p> <p>Lokasi: Florida, USA</p> <p>Tahun 2023</p> <p>Nama Jurnal SURE</p>	<p>Hasil Penelitian:</p> <p>Jurnal ini membahas metode baru untuk meningkatkan keandalan <i>fine-tuning</i> model <i>machine learning</i> dengan menggunakan optimasi berbasis blok. Penelitian ini menunjukkan bahwa pendekatan ini dapat mengurangi variabilitas hasil <i>fine-tuning</i>.</p> <p>Alasan Menjadi Tinjauan Penelitian:</p> <p>Metode yang diusulkan dapat diterapkan pada pengembangan sistem deteksi kekerasan fisik berbasis video untuk meningkatkan keakuratan dan keandalan deteksi khusus nya pada <i>MobileNet</i>.</p>

1.6. Sistematika Penulisan

Laporan Tugas Akhir ini terdiri dari rangkaian sistematika penulisan sebagai berikut:

a) BAB 1 Pendahuluan

Memberikan informasi berupa latar belakang masalah yang diangkat yaitu kekerasan fisik, penelitian sebelumnya, teknologi yang akan digunakan peneliti, rumusan masalah, tujuan penelitian, batasan masalah, dan *State of the Art*.

b) BAB 2 Tinjauan Pustaka

Membahas landasan teori yang menyertai penelitian ini guna memperkuat pokok argumentatif teoritis dalam pengembangan penulisan dan produk.

c) BAB 3 Metode

Membahas metodologi penelitian yang digunakan oleh peneliti sebagai rangkaian penelitian.

d) BAB 4 Pembahasan

Membahas tahapan dalam membangun model deteksi kekerasan fisik berupa pengumpulan data, *pre-processing*, serta analisis performa dan evaluasi model.

e) BAB 5 Kesimpulan dan Saran

Hasil rangkuman dan saran untuk penelitian.

