

# Substraksi Latar Menggunakan Nilai Mean untuk Klasifikasi Kendaraan Bergerak Berbasis Deep Learning

Ilal Mahdi<sup>1</sup>, Kahlil Muchtar<sup>2,3\*</sup>, Fitri Arnia<sup>1,2</sup>, dan Tia Ernita<sup>3</sup>

<sup>1</sup>Program Studi Magister Teknik Elektro, Jurusan Teknik Elektro dan Komputer  
Fakultas Teknik Universitas Syiah Kuala

Jl. Tgk. Syech Abdurrauf No.7 Darussalam, Banda Aceh 23111

<sup>2</sup>Pusat Riset Telematika, Universitas Syiah Kuala

<sup>3</sup>Program Studi Teknik Komputer, Jurusan Teknik Elektro dan Komputer  
Fakultas Teknik Universitas Syiah Kuala

Jl. Tgk. Syech Abdurrauf No.7 Darussalam, Banda Aceh 23111

e-mail: kahlil@unsyiah.ac.id

**Abstrak**—Sistem deteksi objek bergerak telah banyak digunakan dalam kehidupan sehari-hari. Saat ini penelitian dibidang substraksi latar masih terus dilakukan untuk mencapai hasil akurasi yang maksimal. Penelitian ini bertujuan untuk memodelkan substraksi latar dari sebuah citra menggunakan nilai *mean* dengan konsep *non overlapping block*. Selanjutnya, hasil substraksi latar akan digunakan dalam deteksi objek bergerak berbasis *deep learning*. Secara spesifik, citra masukan akan dibagi menjadi beberapa blok, kemudian nilai mean dari setiap blok akan dihitung untuk nantinya menghasilkan blok biner (*binary map*). Blok biner yang telah dihasilkan akan dijadikan sebagai masukan pembangkitan model latar (*background modelling*). Model latar bertujuan untuk memisahkan objek bergerak dengan latar yang ada pada citra masukan. Objek bergerak yang dihasilkan (lokalisasi objek) akan dikirimkan ke tahap klasifikasi objek menggunakan *deep learning*. *Dataset* yang digunakan dalam penelitian ini adalah CDNet 2014. Hasil penelitian mampu menghasilkan sistem deteksi objek bergerak yang lebih akurat. Pengujian kuantitatif yang dilakukan menghasilkan akurasi di atas 90%.

**Kata kunci:** *deteksi objek bergerak, nilai mean, deep learning, substraksi latar*

**Abstract**— Moving object detection systems have been widely used in everyday life. Currently, research in the field of background subtraction is still being carried out to achieve maximum accuracy results. This study aims to model the background subtraction of an image using the mean value with the concept of non overlapping block. Furthermore, the background abstraction results will be used in deep learning-based moving object detection. Specifically, the input image will be divided into several blocks, then the mean value of each block will be calculated to later produce a binary block (binary map). The binary blocks that have been generated will be used as input for background modeling. The background model aims to separate moving objects from the background in the input image. The resulting moving object (object localization) will be sent to the object classification stage using deep learning. The dataset used in this study is CDNet 2014. The results of the study were able to produce a more accurate moving object detection system. Quantitative tests carried out resulted in an accuracy of above 90%.

**Keywords:** *moving object detection, mean value, deep learning, background subtraction*

## I. PENDAHULUAN

Teknik pengolahan citra digital saat ini sudah sangat banyak penerapannya dalam berbagai aplikasi. Teknik ini tidak hanya terbatas digunakan pada proses pengenalan foto atau gambar saja, melainkan juga sudah mulai diterapkan dalam pengenalan gambar bergerak atau biasa disebut video. Di antara pemanfaatan teknik pengolahan citra pada video adalah adanya penggunaan sistem deteksi gerak dan sistem pengawasan objek diam maupun objek yang bergerak. Sistem pengawasan objek di dalam video seringkali diterapkan untuk mengawasi objek tertentu

pada suatu tempat dengan menggantikan peran manusia dalam pengawasan [1].

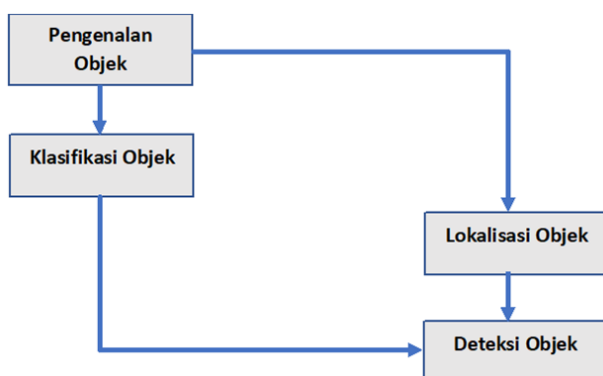
Dalam sistem deteksi objek bergerak, objek tersebut dapat dideteksi dengan cara memanfaatkan informasi dari kumpulan frame pada sebuah video. Metode yang paling umum digunakan dalam pendekatan *spatial-temporal information* adalah metode *frame difference*. Cara kerja metode ini adalah dengan membandingkan kemudian mengurangi latar dari *frame* yang sekarang dengan *frame* yang sudah ada sebelumnya dari beberapa urutan *frame* dari sebuah video. Apabila selisih antara piksel sekarang dengan piksel dari *frame* sebelumnya bernilai lebih besar

dari suatu nilai ambang yang dipilih, maka objek pada *frame* tersebut dianggap sebagai objek yang bergerak.

Dalam proses pengenalan objek yang bergerak, metode deteksi dan klasifikasi objek bergerak yang sangat familiar adalah metode berbasis *deep learning*. Di antara kekurangan metode konvensional adalah proses komputasi yang besar dan sulit dijalankan terutama dalam beberapa kasus yang kompleks seperti pada kasus oklusi di antara kendaraan. *You Only Look Once (YOLO) version 4* merupakan metode yang dibangun untuk mendeteksi objek dalam sebuah citra dengan keunggulan kecepatan proses dan akurasi yang tinggi [2]. YOLO dibangun menggunakan arsitektur *deep learning* DarkNet-53 sebagai model dasar.

Dalam penerapannya, deteksi objek bergerak menggunakan *deep learning* masih mengalami kendala dengan adanya kondisi sebuah *frame* yang tidak ideal. Selain itu, kondisi *frame* yang mengalami berbagai permasalahan seperti pencahayaan yang tidak merata, adanya benda lain yang ikut bergerak ketika pengambilan *frame*, dan tingkat kontras yang rendah dapat menyebabkan akurasi dari sistem deteksi objek menurun. Oleh karena itu, beberapa peneliti menggunakan teknik penghapusan latar untuk meningkatkan kembali akurasi dari deteksi objek menggunakan *deep learning* seperti [1-4]. Kinerja metode substraksi latar sendiri bergantung pada teknik pemodelan yang digunakan. Hal ini karena latar belakang pada video jauh lebih dinamis dibandingkan latar belakang gambar dan biasanya bergerak. Beberapa metode substraksi latar akan turut mengasumsikan pergerakan latar belakang ini sebagai *Region of Interest (ROI)* atau area objek. Hal ini terjadi karena kebanyakan metode substraksi latar hanya menggunakan piksel warna atau informasi intensitas piksel untuk melakukan segmentasi citra seperti algoritma *Gaussian Mixture Model (GMM)* [7].

Pada penelitian lainnya, metode yang dipakai dalam proses deteksi objek bergerak adalah metode penghapusan latar berbasis blok adaptif, metode tersebut terinspirasi dari blok berdasarkan *Adaptive Segmenter* [3]. Penelitian sebelumnya menunjukkan bahwa sebuah perangkat untuk mendeteksi objek memerlukan sebuah pemodelan latar yang kompleks dan proses untuk menghadapi perubahan iluminasi, maka penulis mengusulkan pendekatan adaptif berdasarkan representasi tekstur blok yang tidak tumpang



Gambar 1. Penelitian dalam bidang visi komputer [10]

tindih, yang tujuannya untuk merancang cahaya komputasi dan solusi yang efisien untuk meningkatkan ketahanan deteksi.

Pada penelitian ini, penulis mengusulkan teknik substraksi latar berbasis fitur tekstur dan algoritma *Bit Truncation Coding (BTC)* untuk lokalisasi dan deteksi objek mobil. Fitur tekstur dimanfaatkan karena menunjukkan kinerja yang memuaskan dalam membedakan informasi dalam sebuah citra [4], [5]. Fitur tekstur berfungsi untuk mendapatkan informasi dalam sebuah *frame* berdasarkan perbedaan intensitas dalam *frame/citra*. Pemanfaatan algoritma BTC dalam proses substraksi latar membuat proses substraksi latar lebih terfokus pada area tertentu. Hal ini disebabkan BTC memproses sebuah *frame/citra* secara *non-overlapping block* (blok atau area yang tidak tumpang tindih). Secara spesifik, algoritma BTC membagi setiap *frame/citra* menjadi beberapa blok yang berukuran sama untuk setiap bloknya dan setiap blok tidak saling tumpang tindih (*non-overlapping block*). Setiap nilai-nilai piksel yang ada pada setiap blok dilakukan perhitungan nilai rata-rata (*mean*). Dibandingkan dengan algoritma yang memanfaatkan blok tumpang tindih, BTC mampu memproses tiap masukan lebih cepat (dari sisi *frame per second/FPS*). Walaupun tekstur yang dihasilkan lebih kasar (*coarse*), objek bergerak telah dapat disubtraksi dari latarnya secara akurat [1]. Sebelum diterapkan teknik substraksi latar, terlebih dahulu dilakukan proses penekanan *noise* menggunakan metode *guided filter*. Selain itu, pada penelitian ini juga diterapkan operator morfologi citra (*erosi dan dilasi*) untuk mengekstraksi komponen citra.

## II. STUDI PUSTAKA

### A. Deteksi dan Klasifikasi Objek Konvensional

Gambar 1 menunjukkan beberapa teknik dalam bidang visi komputer untuk mengekstraksi objek pada sebuah citra digital. Jika dilihat secara mendalam terdapat tiga bagian dari bidang visi komputer yang terlihat sangat mirip, namun punya kajian yang berbeda, yaitu: klasifikasi citra, lokalisasi objek, dan deteksi objek. Dalam implementasinya, terkadang pengenalan objek sering dipahami sebagai deteksi objek [9].

Dalam proses deteksi objek bergerak terdapat beberapa proses pengolahan citra dan pembelajaran mesin (*machine learning*) yang dilakukan sehingga menghasilkan suatu teknik deteksi objek yang bergerak. Kemudian dilakukan segmentasi terhadap citra atau *frame* masukan, lalu



Gambar 2. Objek atau kendaraan bergerak [12]

dilanjutkan dengan lokalisasi dan terakhir adalah klasifikasi objek yang menjadi target deteksi.

Berdasarkan Gambar 2, analisis awal dari proses deteksi objek biasanya dimulai ketika ada objek yang bergerak atau berpindah di dalam sebuah video. Sistem yang bekerja untuk mendeteksi setiap pergerakan objek yang terjadi di dalam video disebut sistem deteksi gerakan. Metode yang paling umum digunakan dalam pendekatan *spatial-temporal information* adalah metode *frame difference*. Metode ini akan melakukan perhitungan menyeluruh pada nilai piksel dari setiap *frame*. Apabila selisih antara piksel suatu *frame* dengan piksel dari *frame* sebelumnya bernilai lebih besar dari suatu nilai ambang yang dipilih, maka objek pada *frame* tersebut dianggap sebagai objek yang bergerak.

### B. Deteksi dan Klasifikasi Objek melalui Deep Learning

*Deep learning* adalah suatu teknik pembelajaran yang diberikan kepada mesin untuk belajar dengan jumlah data yang sangat banyak [10]. Pada Gambar 3 diilustrasikan sebuah struktur sederhana *Convolutional Neural Network* (CNN) yang umum digunakan. Pada CNN, tahapan pra-pemrosesan yang digunakan relatif sedikit dibandingkan dengan metode klasifikasi lainnya. Hal ini berarti bahwa CNN belajar menggunakan *filter* yang lebih dalam dibandingkan algoritma pembelajaran mesin yang tradisional yang menggunakan teknik *handcrafted*. Penggabungan kemampuan untuk belajar sendiri dan campur tangan manusia dalam mendesain fitur merupakan kelebihan utama dalam mengaplikasikan CNN.

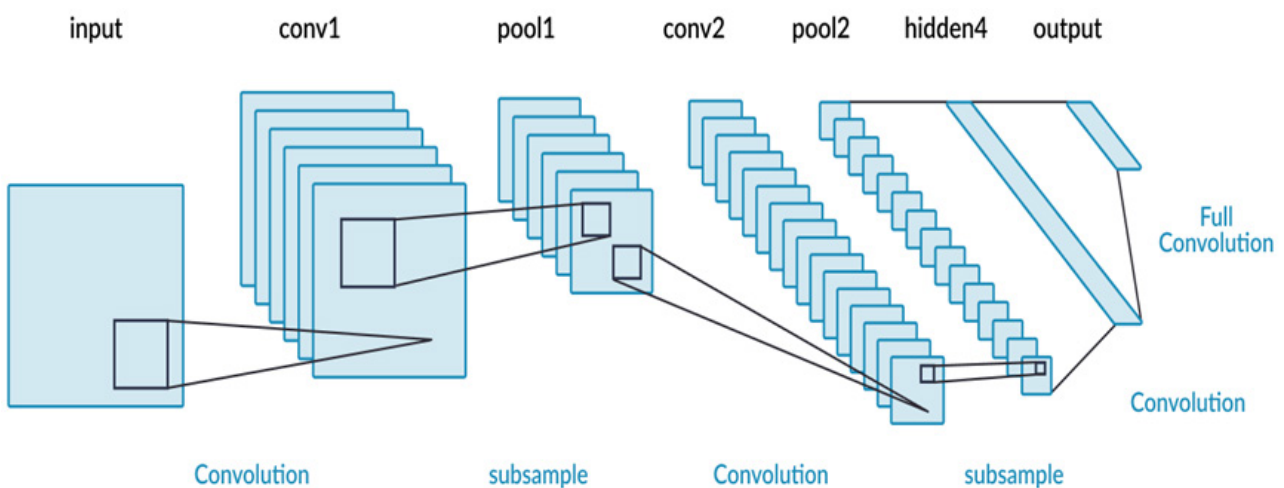
*You Look Only Once* (YOLO) merupakan salah satu model *deep learning* yang dapat digunakan untuk pengenalan objek. YOLO adalah serangkaian model pembelajaran mendalam *end-to-end* yang dirancang untuk deteksi objek secara cepat, yang dikembangkan oleh Joseph Redmon, Santosh Divvala, Ross Girshick, dan Ali Farhadi pada tahun 2015 dalam artikel mereka yang berjudul “*You Look Only Once: Unified, Real-Time Object Detection*” [2]. YOLO menerapkan *neural network*

pada sebuah citra, kemudian membagi citra menjadi daerah dan memprediksi *bounding box* serta probabilitas untuk masing-masing daerah *bounding box*. Probabilitas untuk setiap *bounding box* kemudian dihitung untuk mengklasifikasikan sebagai objek atau bukan. YOLO dapat melakukan pengenalan objek secara *real-time* dengan kecepatan 45 *frame per second*. Pendekatan ini melibatkan jaringan saraf *convolutional* tunggal yang dalam pendeteksiannya membagi *input* ke dalam kotak sel dan setiap sel secara langsung memprediksi *bounding box* dan klasifikasi objek. Secara spesifik, Gambar 4 selanjutnya menjelaskan bahwa arsitektur dasar dari YOLOv4 adalah DarkNet-53. DarkNet-53 adalah sebuah model klasifikasi CNN yang dibangun atas fondasi ResNet [2].

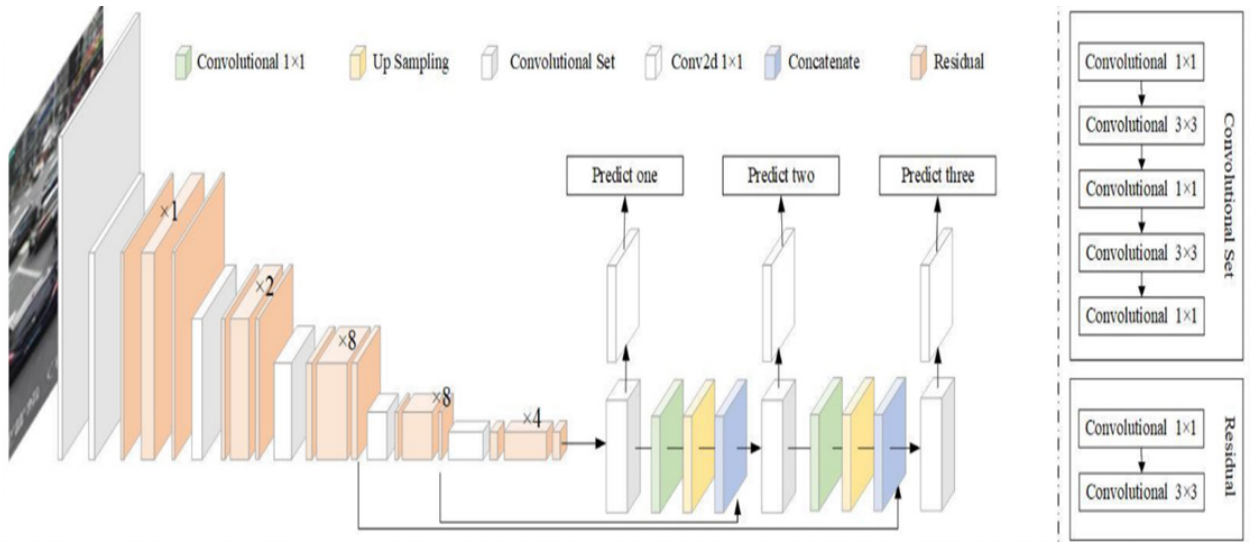
### III. METODE/DESAIN

Gambar 5 menunjukkan tahapan klasifikasi kendaraan bergerak secara komprehensif. Tahapan utama adalah melakukan penggabungan metode substraksi latar dan *deep learning* untuk mendapatkan hasil akhir yang akurat dan efisien.

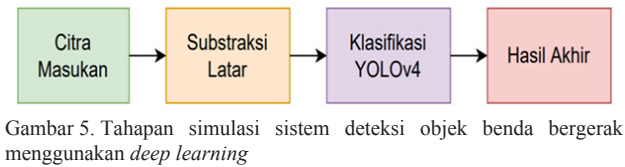
Seperti yang ditunjukkan pada Gambar 6, secara spesifik, tahapan pertama dalam proses substraksi latar adalah akuisisi video masukan berupa hasil rekaman CCTV yang dipasang di jalan raya. Hasil akuisisi video disimpan dalam format MP4. Selanjutnya, dilakukan proses ekstraksi *frame* dari video masukan menjadi format *joint photographic experts group* (JPEG). Pada tahapan ekstraksi *frame* ini, 500 *frame* pertama akan diasumsikan sebagai *background model*. *Background model* adalah kumpulan *frame* awal yang hanya berisi latar belakang tanpa adanya objek yang bergerak yaitu mobil. Sedangkan *frame* selain 500 *frame* awal dijadikan sebagai target deteksi untuk melihat keberadaan mobil yang bergerak. Tahapan berikutnya adalah *denoising* atau penekanan *noise* menggunakan teknik *guided filter* untuk menghasilkan citra masukan yang lebih lembut (*smooth*). Selanjutnya, Algoritma BTC diaplikasikan untuk membagi setiap *frame*/citra menjadi beberapa blok yang berukuran sama



Gambar 3. Struktur sederhana CNN berbasis *deep learning* [11]



Gambar 4. Arsitektur dari DarkNet-53 yang digunakan pada YOLO [2]



Gambar 5. Tahapan simulasi sistem deteksi objek benda bergerak menggunakan deep learning

dan tidak saling tumpang tindih.

Setelah didapatkan nilai rata-rata dari setiap blok, algoritma substraksi latar akan membandingkan blok latar (*background*) yang didapatkan dari ekstraksi 500 *frame* awal dengan *frame* berikutnya. Hasil substraksi menghasilkan selisih nilai piksel untuk memastikan keberadaan mobil sebagai objek yang bergerak. Jika keberadaan objek tidak ditemukan, maka hasil substraksi mengembalikan *frame* yang kosong.



Gambar 6. Alur teknik penghapusan latar



Gambar 7. Teknik *non-overlapping* blok pada BTC

#### A. Block Truncation Coding (BTC)

BTC pertama kali diusulkan untuk mengembangkan algoritma kompresi citra. Untuk mendeteksi objek bergerak, metode BTC bekerja dengan membagi sebuah citra ke dalam blok 4x4 dengan bentuk dan ukuran yang sama. Setiap blok pada citra memiliki nilai rata-rata piksel (*mean*) yang digunakan untuk mengubah blok menjadi blok biner. Jika *mean* dari suatu blok lebih besar dari nilai piksel yang sedang diobservasi, maka nilai piksel tersebut diubah menjadi 1. Sebaliknya, jika *mean* dari suatu blok lebih kecil dari nilai piksel maka piksel diubah menjadi 0 [4]. Ilustrasi pembangkitan blok BTC dapat dilihat pada Gambar 7:

Sebagaimana dijelaskan pada [15], ukuran blok 4x4 mampu menghasilkan informasi tekstur yang paling baik, jika dibandingkan dengan ukuran blok lain, yaitu ukuran 2x2 dan 6x6. Pada perbandingan tersebut, ukuran blok 2x2 menghilangkan terlalu banyak informasi tekstur, sedang blok 6x6 menampilkan detail yang tidak penting pada blok biner yang dihasilkan.

#### B. Guided Filter sebagai Pra-proses

*Guided filter* merupakan filter *smoothing* yang dipakai sebagai teknik *pre-processing* untuk menghilangkan derau/*noise* pada citra dengan tetap menjaga bentuk dasar dari tepian objek pada citra tersebut. Tujuannya adalah agar informasi citra yang ada tidak hilang. Ilustrasi proses

*guided filter* dan filter bilateral dapat dilihat pada Gambar 8.

Seperti yang terlihat pada gambar tersebut, proses komputasi dari *bilateral filter* lebih kompleks daripada proses komputasi *guided filter* yang linier. *Guided filter* dalam perhitungannya hanya membutuhkan input atau masukan  $p$  dan *guidance image*  $I$ . Citra output filtering  $q$  yang dihasilkan dari proses filtering dapat mempertahankan informasi utama dari gambar input dan cenderung mendapatkan variasi dari *guidance image*.

C. Klasifikasi Objek berbasis YOLO

Penelitian ini dilakukan pada komputer dengan sistem operasi Windows 10, menggunakan *software* Visual Studio Code Editor, OpenCV beserta *library-library* yang dibutuhkan. Pengujian algoritma menggunakan kumpulan *frame* dari *website* penyedia *dataset* yaitu CDNET. *Dataset* yang dipilih adalah *dataset highway* dan *wetSnow*, dapat dilihat pada Gambar 9 dan Gambar 10. Selain itu digunakan pemodelan *deep learning* YOLOv4 sebagai metode deteksi. Metode YOLO merupakan metode yang efektif untuk mendeteksi objek berdasarkan regresi daripada melakukan klasifikasi.

1. Konfigurasi Dataset

*Dataset* yang digunakan merupakan *dataset* yang

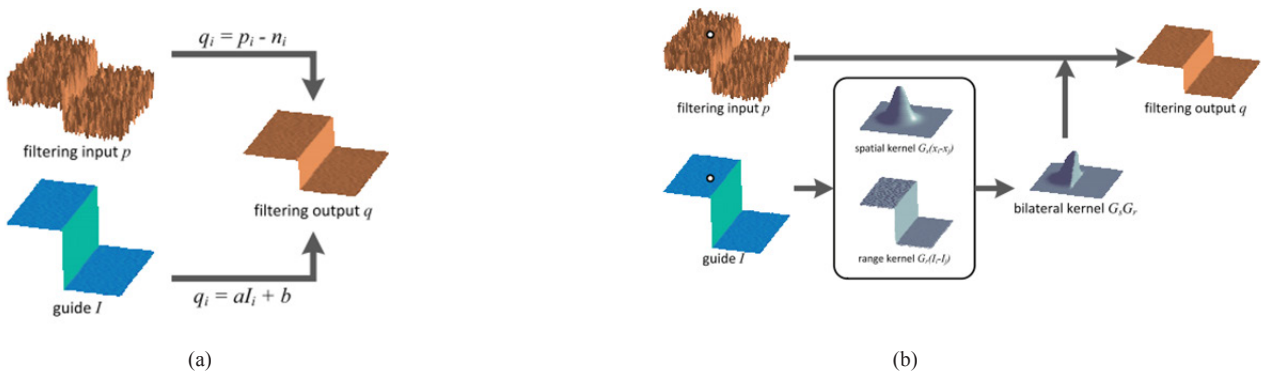
disediakan oleh CDnet 2014. Citra masukan adalah kumpulan *frame* yang berasal dari video yang sudah diekstraksi *frame-nya*. Proses penentuan piksel dilakukan sesudah membagi citra menjadi blok area dengan ukuran 4x4 piksel. Kemudian *frame* citra hasil penghapusan latar dimasukkan ke dalam metode deteksi dan klasifikasi objek bergerak, yaitu metode *deep learning* menggunakan model YOLOv4.

Hasil substraksi latar selanjutnya dijadikan masukan pada deteksi objek menggunakan YOLOv4. Hasil deteksi menggunakan YOLOv4 sesudah di substraksi latar dibandingkan dengan hasil deteksi tanpa substraksi latar.

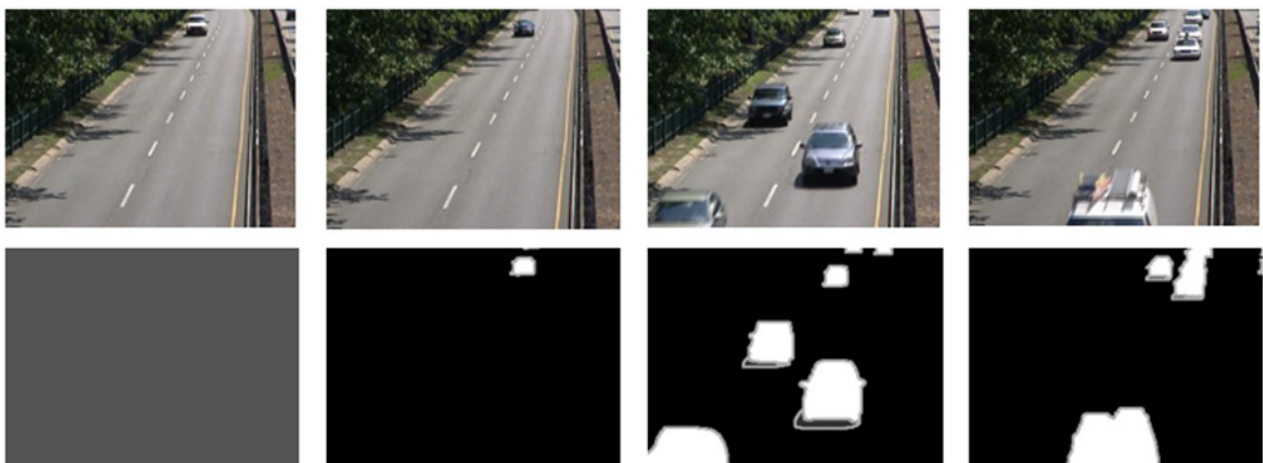
Citra masukan sesudah proses penghapusan latar dimasukkan ke metode klasifikasi YOLOv4. YOLOv4 melakukan proses deteksi melalui empat tahapan atau langkah yaitu tahapan yang dinamakan *backbone*, kemudian tahapan *Neck*, Tahapan *Dense prediction*, dan terakhir tahapan *sparse prediction*. Keluaran dari tahapan *sparse prediction* terlihat seperti pada bagian *output* pada Gambar 11.

2. Prosedur Pengujian

Proses evaluasi akurasi algoritma untuk substraksi latar dapat dilakukan dengan berbagai macam metode. Metode yang digunakan untuk mengetahui dan membuktikan keakuratan sebuah algoritma disebut dengan evaluasi



Gambar 8. Ilustrasi proses penapisan menggunakan: (a). Guided Filter; (b). Filter Bilateral [13]



Gambar 9. Dataset highway dari CDnet 2014 [11]

Tabel 1. *Confusion Matrix*

Acuan/ Pengajuan	Foreground	Background
Foreground	TP	FN
Background	FP	TN

performa (*performance evaluation*). Hasil perbandingan piksel-piksel berupa nilai dapat diklasifikasikan dalam sebuah matriks yang disebut *confusion matrix* [14] seperti ditunjukkan pada Tabel 1.

Parameter pengukuran keakuratan metode yang akan digunakan dalam penelitian ini adalah *Precision* (Pr), *Recall* (Re), *Specificity* (Sp), dan *Accuracy* (Acc).

Selain itu, untuk memastikan keakuratan metode yang diusulkan, dilakukan perbandingan nilai pengujian kuantitatif yang telah diperoleh dengan nilai dari GMM. Nilai hasil pengujian kuantitatif dari metode GMM dapat diperoleh dari perhitungan dengan metode pengujian kuantitatif yang sama.

$$Precision(Pr) = \frac{TP}{TP + FP}, \tag{1}$$

$$Recall(Re) = \frac{TP}{TP + FN}, \tag{2}$$

$$Specificity(Sp) = \frac{TN}{TN + FP}, \tag{3}$$

Tabel 2. Contoh perhitungan *frame rates*

Jumlah Frame Dataset X	2700 frames
Waktu Pemrosesan Algoritma Y	4 menit 13 detik (253 second)

$$Accuracy(Acc) = \frac{TP}{P + N}. \tag{4}$$

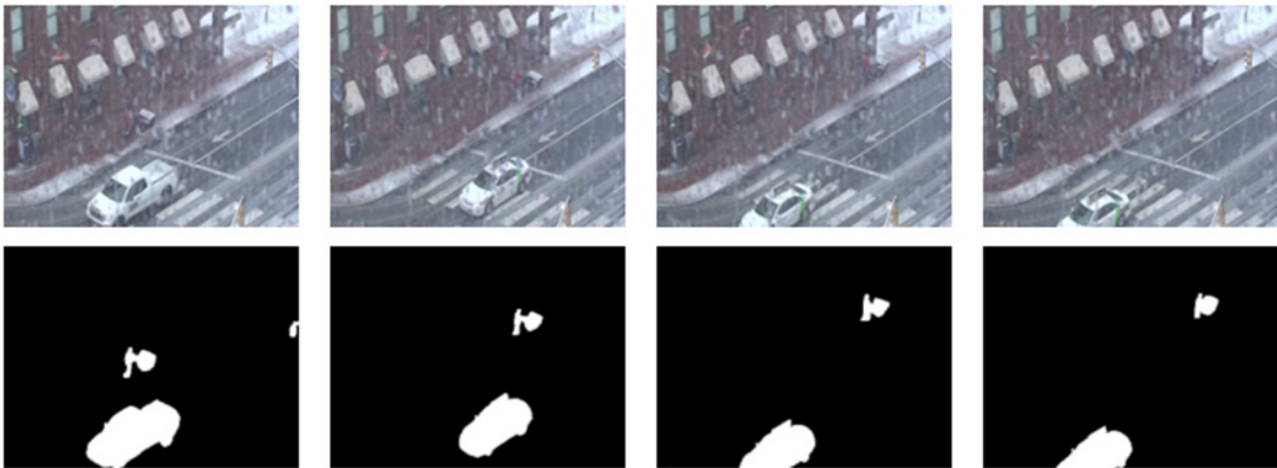
Selanjutnya, contoh sederhana perhitungan *frame rate* adalah sebagai berikut:

$$Frame\ Rates = \frac{2700\ frame}{253\ second} = 10,6\ fps.$$

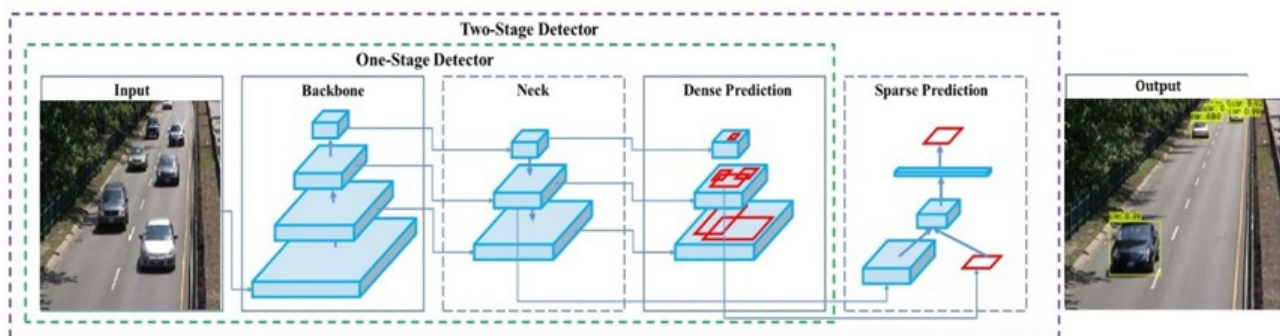
Berdasarkan perhitungan, algoritma Y yang menggunakan *dataset X* sebagai masukan memiliki *frame rates* sebesar 10,6 fps. Artinya adalah ada ~10 *frame dataset X* setiap detiknya yang berhasil diproses oleh algoritma Y.

#### IV. HASIL DAN PEMBAHASAN

Penelitian ini menghasilkan sebuah sistem deteksi objek bergerak berbasis deep learning dengan memanfaatkan lokalisasi benda bergerak. Sistem ini memanfaatkan teknik preprocessing berupa metode penekanan *noise* (*denoising*) menggunakan *guided filter*. Kemudian dilanjutkan dengan proses substraksi latar dan deteksi objek menggunakan YOLOv4.



Gambar 10. *Dataset wetSnow* dari CDnet 2014 [11]



Gambar 11. Klasifikasi objek bergerak menggunakan YOLOv4

### A. Hasil Substraksi Latar

Gambar 12 mengilustrasikan perbandingan secara kualitatif antara metode yang diusulkan dengan metode tekstur sebelumnya [5]. Metode yang diusulkan oleh Heikkilä dan Pietikäinen [5] menggunakan blok yang tumpang-tindih (*overlapping*) sehingga membutuhkan waktu proses yang cukup lama (menggunakan CPU). Oleh karena itu, walaupun metode yang diusulkan menghasilkan luaran yang lebih kasar (*coarser*), namun tetap dapat digunakan sebagai pra-proses penentuan ROI yang akurat dan efisien. Pada Gambar 12 digambarkan seseorang mendekati kamera dan disaat yang sama proses substraksi latar dilakukan secara *real-time*. Tampak jelas bahwa metode yang diusulkan mampu memisahkan derau (*noise*) dari objek bergerak.

Gambar 13 dan Gambar 14 menunjukkan hasil lokalisasi objek mobil pada *dataset highway (high)* dan *wetSnow (WS)* menggunakan metode substraksi latar yang digabungkan dengan metode *guided filtering* dan operator morfologi. Secara berurutan, tahapan lokalisasi objek mobil diawali dengan proses *smoothing* pada citra masukan menggunakan metode *guided filtering*. Selanjutnya dilakukan proses substraksi latar untuk memisahkan objek mobil dan latar belakang. Tahap

terakhir yaitu menggunakan operator morfologi untuk memperbaiki bentuk objek.

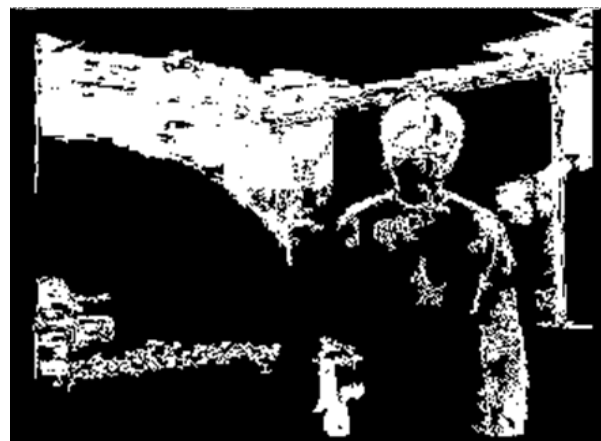
### B. Perbandingan Metode Substraksi Latar dengan GMM

Gambar 15 menunjukkan perbandingan kinerja metode substraksi latar untuk lokalisasi kendaraan bergerak antara metode usulan dan GMM. Kinerja hasil dibandingkan berdasarkan kemiripan secara visual antara citra hasil dengan citra GT. Perbandingan dilakukan terhadap citra hasil substraksi latar.

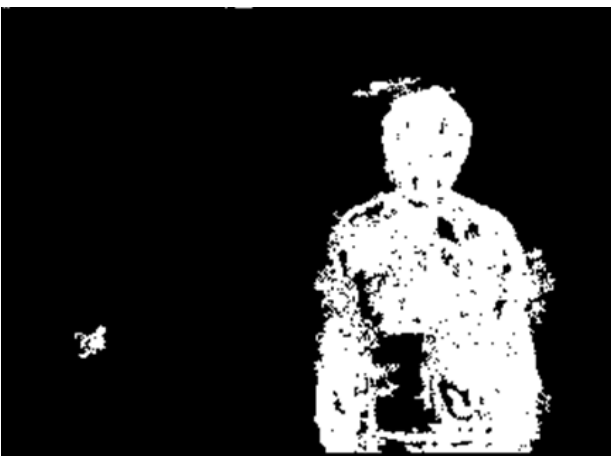
Gambar 15 menunjukkan perbandingan hasil deteksi objek mobil menggunakan metode yang diusulkan, metode GMM, dan *ground truth* (gambar acuan) dari *frame* yang dipilih. Hasil pengujian ditampilkan dalam bentuk gambar biner dimana objek yang bergerak (mobil) diwakili oleh warna putih sedangkan latar belakang diwakili oleh warna hitam. Terlihat bahwa metode usulan maupun metode GMM mampu mendeteksi objek mobil sesuai dengan gambar acuan pada masing-masing *frame*. Perbedaan signifikan yang terlihat adalah objek mobil yang terdeteksi pada metode usulan memiliki lebih banyak lubang seperti yang terlihat pada *frame* 817 dan 1668 dari *dataset highway* dan *frame* 512 dan 738 dari *dataset wetSnow*. Namun, berdasarkan *frame* 512 dan 738 dari *dataset*



(a)



(b)



(c)



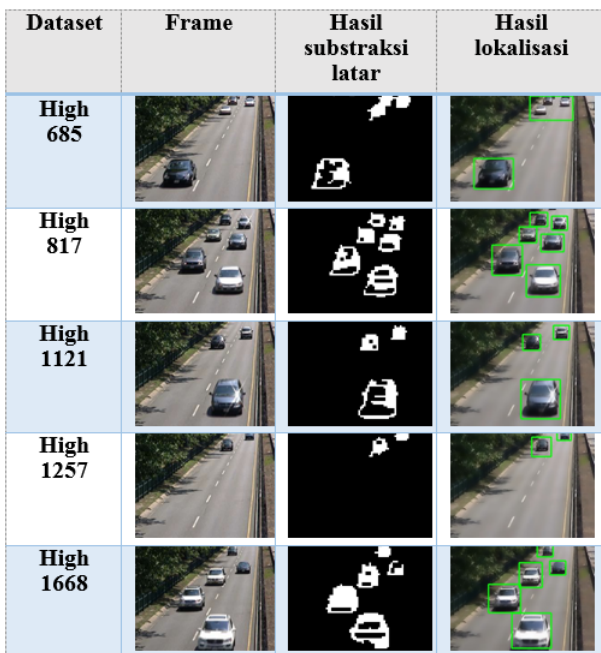
(d)

Gambar 12. Perbandingan hasil kualitatif antara (b) GMM [7], (c) metode Heikkilä dan Pietikäinen [5], dan (d) metode yang diusulkan

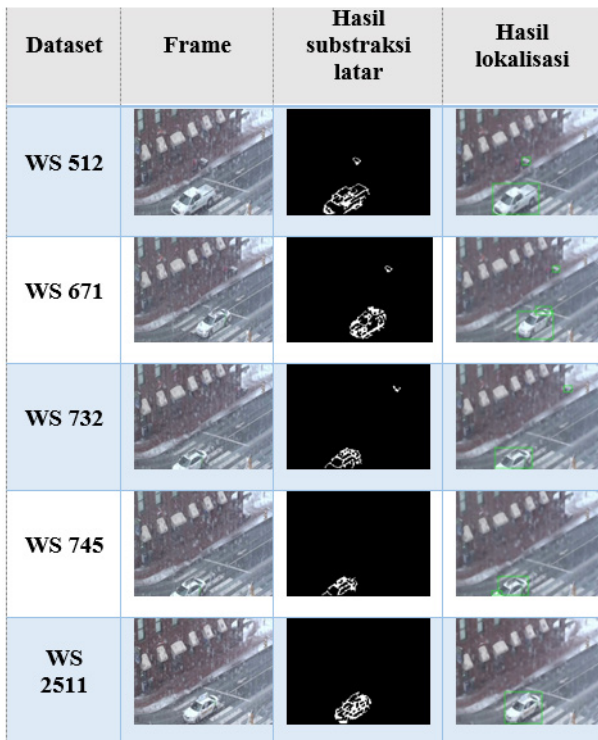
wetSnow terlihat bahwa hanya terdapat sedikit *noise* pada hasil deteksi menggunakan metode usulan. Sedangkan pada *frame* yang sama dengan metode GMM, terdapat banyak *noise* yang ikut dianggap sebagai objek.

C. Perbandingan Hasil Pengujian secara Kuantitatif

Tabel 3 menunjukkan perbandingan hasil pengujian secara kuantitatif antara metode usulan dengan metode



Gambar 13. Hasil substraksi latar dan lokalisasi objek mobil



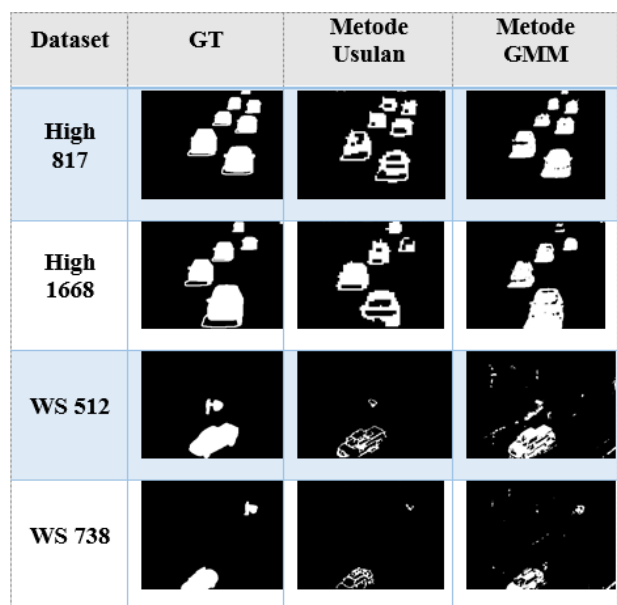
Gambar 14. Hasil substraksi latar dan lokalisasi objek mobil pada dataset wetSnow

GMM. Dalam mendeteksi objek mobil pada *dataset highway*, setiap parameter hasil pengujian yaitu *precision*, *specificity*, *recall*, dan *accuracy* dari metode yang diusulkan memiliki nilai yang mendekati nilai pengujian dari metode GMM. Hal ini menunjukkan bahwa metode yang diusulkan merupakan metode yang akurat dalam mendeteksi objek mobil saat kondisi pencahayaan *frame* baik meskipun terdapat banyak *noise* berupa pergerakan daun.

Berdasarkan tabel 3, hasil evaluasi terhadap *dataset* dengan kategori *wetSnow* menunjukkan hasil yang berbeda dari *dataset* kategori *highway*. Metode usulan menunjukkan kinerja maksimal atau hampir mendekati 100% dalam parameter *specificity*. Begitu pula untuk

Tabel 3. Perbandingan hasil pengujian kuantitatif dari algoritma yang diusulkan dan GMM

Dataset	Metode	Parameter (%)			
		Precision	Specificity	Recall	Accuracy
Frame 817 "Highway"	Metode Substraksi Latar	75,74	96,60	63,25	91,81
	GMM	88,431	98,26	73,14	95,52
Frame 1668 "Highway"	Metode Substraksi Latar	74,38	95,98	62,04	90,6
	GMM	82,92	97,06	75,82	93,69
Frame 512 "WetSnow"	Metode Substraksi Latar	73,73	99,33	25,6	94,27
	GMM	70,63	97,93	67,55	95,85
Frame 738 "WetSnow"	Metode Substraksi Latar	93,67	99,91	30,92	97,05
	GMM	85,29	99,55	60,3	97,92



Gambar 15. Perbandingan metode usulan dengan GMM dalam mensubstraksi latar pada objek bergerak



hasil pengujian dengan parameter *precision* dan *accuracy*. Bahkan nilai *precision* pada *frame* 738 dengan metode usulan memiliki selisih sampai 8,38% dengan metode GMM. Namun, untuk parameter *recall* hasil pengujian menggunakan metode yang diusulkan mengalami banyak penurunan sampai lebih dari 30% dibandingkan dengan metode GMM. Hal ini menunjukkan bahwa metode yang diusulkan mampu mendeteksi objek serta mengurangi *noise* dengan baik, namun sulit mendefinisikan bentuk objek pada kondisi *frame* dengan *noise* yang lebih kompleks. Berdasarkan Gambar 12 dan 13, terlihat bahwa kondisi citra pada *dataset wetSnow* mempunyai kondisi yang lebih buruk dari *dataset highway*.

Gambar 16 menunjukkan nilai akurasi tertinggi adalah 1,00 yang artinya metode YOLOv4 yakin 100% bahwa objek tersebut adalah mobil, sedangkan nilai terendah adalah 0,92 yang artinya kemungkinan objek tersebut berupa mobil adalah 92%. Pada Gambar 15c, nilai deteksi objek tertinggi adalah 0,99 yang berarti 99% objek tersebut dikenali sebagai mobil. Nilai terendah adalah 0,95 yang berarti probabilitas objek tersebut berupa mobil adalah 95%. Hal ini disebabkan oleh jumlah mobil yang ada dalam *frame* pada Gambar 15(c) yang hanya terdiri dari jumlah mobil yang sedikit.

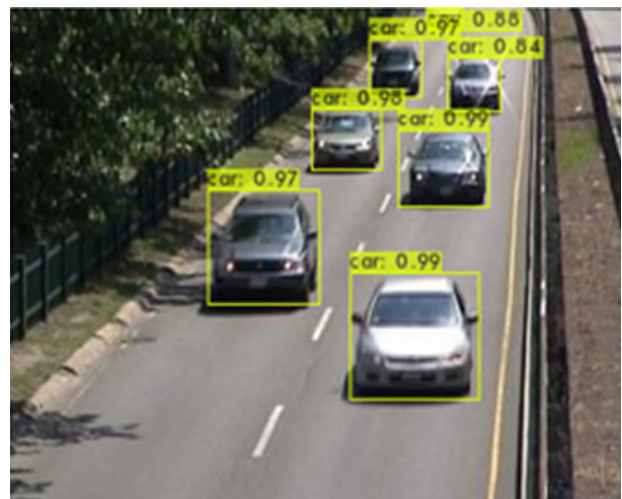
Tabel 4 menampilkan perbandingan waktu pengolahan (*processing time*) dari beberapa metode terdahulu dan metode yang diusulkan. Terlihat bahwa jumlah *frame* (*frame rates*) yang dapat diproses oleh setiap metode dalam satuan waktu *second* (*fps*) sangat beragam. Berdasarkan data hasil pengujian, metode yang diusulkan dapat memproses sampai 55 *fps* sedangkan GMM hanya mampu memproses 21 *fps*. Hal ini menunjukkan bahwa metode usulan lebih

Tabel 4. Perbandingan FPS antar metode

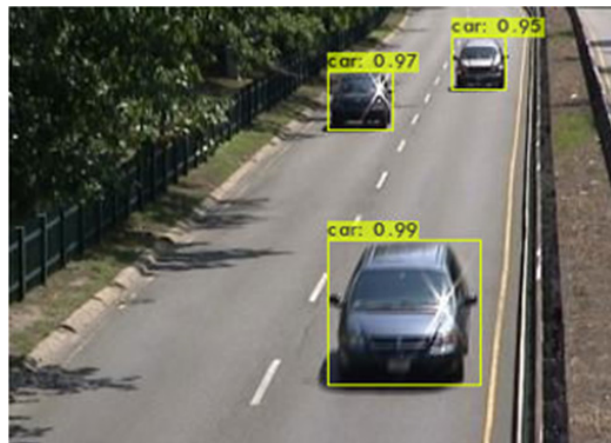
Metode	Hasil FPS
Semantic Background Substraction	7
IUTIS-3	10
WisenetMD	12
PAWCS	27
WeSamBE	2
SuBSENSE	30
GMM	21
Improved GMM	49
Substraksi Latar Berbasis Tekstur (metode usulan)	55



(a)



(b)



(c)

Gambar 16. Akurasi pengenalan mobil sesudah proses lokalisasi, dimana: (a) *frame* 685, (b) *frame* 817, (c) *frame* 1121

unggul dalam hal efisiensi waktu pengolahan. Penggunaan fitur tekstur dan BTC memungkinkan metode yang diusulkan memproses lebih banyak *frame* setiap detiknya dibandingkan dengan algoritma pemrosesan berbasis piksel yang secara teori lebih lambat. Kecepatan waktu pengolahan sesuai untuk penggunaan pada perangkat *real-time* seperti untuk mendukung pengembangan *smart city* dalam bidang keamanan dan pengawasan (*surveillance*).

## V. KESIMPULAN

Penelitian ini telah berhasil mengembangkan suatu metode untuk melakukan lokalisasi dan mendeteksi objek bergerak khususnya mobil menggunakan algoritma *Bit Truncation Coding (BTC)*, fitur tekstur dan *guided filtering*. Metode yang diusulkan mampu meminimalkan keberadaan *noise* sekaligus menandai objek mobil dengan baik. Hasil akurasi yang dicapai oleh teknik ini dalam mendeteksi objek mobil mencapai 0,98 atau 98% jika *frame* yang diproses memiliki kondisi pencahayaan baik. Sebaliknya, dengan nilai *recall* terendah yaitu 25,6% pada *frame* 512 *dataset wetSnow*, menunjukkan bahwa metode yang diusulkan masih sulit untuk mendefinisikan bentuk objek jika *frame* blur dan kompleks. Namun begitu, metode yang diusulkan terbukti lebih unggul dalam hal waktu pengolahannya (*processing time*) karena dapat memproses lebih banyak *frame* setiap detiknya dibandingkan metode terdahulu, yaitu sampai 55 fps.

## REFERENSI

- [1] C.H. Yeh, C.Y. Lin, K. Muchtar, and L.W. Kang, "Real-time background modeling based on a multi-level texture description", *Information Sciences*, vol. 269, pp.106-127, 2014.
- [2] A. Farhadi., and J. Redmon., YOLOv4: An incremental improvement. In *Computer Vision and Pattern Recognition*, Berlin/Heidelberg, Germany: Springer, pp. 1804-2767, 2018.
- [3] K. Muchtar, F. Rahman, T.W. Cenggoro, A. Budiarto, and B. Pardamean, "An improved version of texture-based foreground segmentation: block-based adaptive segmenter", *Procedia Computer Science*, vol. 135, pp.579-586, 2018.
- [4] S. Y. Irianto, *Analisa Citra Digital dan Content Based Image Retrieval*, 1st Ed, Bandar Lampung, Indonesia: CV. Anugrah Utama Raharja (AURA), 2016.
- [5] M. Heikkila dan M. Pietikainen, "A Texture-based method for modeling the background and detection of moving objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 657-661, 2006.
- [6] K. He, J. Sun dan X. Tang, "Guided image filtering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1-14, 2013.
- [7] C. Stauffer & W. E. L. Grimson, Adaptive background mixture models for real-time tracking. In *Proceedings. 1999 IEEE computer society conference on computer vision and pattern recognition (Cat. No PR00149)*, Vol. 2, 1999. pp. 246-252.
- [8] A. Susanto, "Penerapan operasi morfologi matematika citra digital untuk ekstraksi area plat nomor kendaraan bermotor," *Jurnal Pseudocode*, vol. 6, no. 1, pp. 49-57, 2019.
- [9] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, and A.C. Berg, "Imagenet large scale visual recognition challenge", *International Journal of Computer Vision*, vol. 115, no. 3, pp.211-252, 2015.
- [10] <https://machinelearningmastery.com/object-recognition-with-deep-learning/>
- [11] Y. Wang, P.M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, and P. Ishwar, P. "CDnet 2014: An expanded change detection benchmark dataset", In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops, USA, 2014*, pp. 387-394.
- [12] J. Redmon, S. Divvala, R. Girshick & A. Farhadi, "You only look once: Unified, real-time object detection" In *Proceedings of the IEEE conference on computer vision and pattern recognition, 2016*, pp. 779-788.
- [13] K. He, J. Sun, and X. Tang, Guided image filtering. In *European conference on computer vision, 2010*, pp. 1-14. Springer, Berlin, Heidelberg.
- [14] N. Nasaruddin, K. Muchtar, and A. Afdhal, "A Lightweight moving vehicle classification system through attention-based method and deep learning," *IEEE Access*, vol. 7, pp. 157564-157573, 2019.
- [15] N. Nasaruddin, K. Muchtar, A. Afdhal, and A. P. J. Dwiyanoro, Deep anomaly detection through visual attention in surveillance videos. *Journal of Big Data*, vol.7, no.1, pp.1-17, 2020.