

Deteksi Kapal di Laut Indonesia Menggunakan YOLOv3

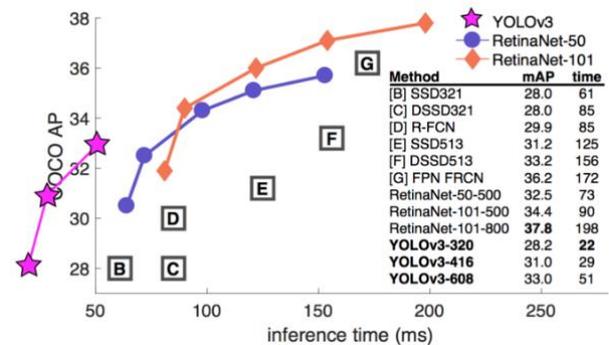
Adam Fahmi Fandisyah, Nur Iriawan dan Wiwiek Setya Winahju
Departemen Statistika, Institut Teknologi Sepuluh Nopember (ITS)
e-mail: nur_i@statistika.its.ac.id

Abstrak—Indonesia adalah negara kepulauan terbesar di dunia yang memiliki kandungan kekayaan dan sumber daya alam laut yang sangat berlimpah. Hal ini memicu terjadinya peristiwa seperti *illegal fishing*, *illegal mining*, *illegal logging*, *drugs trafficking* dan *people smuggling* yang menunjukkan bahwa kurang maksimalnya pengawasan wilayah laut Indonesia. Pesatnya perkembangan teknologi di bidang kecerdasan buatan mendorong ditemukannya *deep learning*, salah satunya yaitu metode *You Only Look Once (YOLO)* yang dikembangkan dengan algoritma untuk mendeteksi sebuah objek secara *realtime*. Dalam penelitian ini, deteksi tipe kapal dilakukan dengan menggunakan YOLOv3 dan dievaluasi dengan menghitung nilai *Mean Average Precision (mAP)* yang dibandingkan hasilnya dengan *ground truth*. Hasil deteksi tipe kapal menggunakan YOLOv3 dengan *k-means anchor box* dapat mengenali tipe kapal pada citra satelit, diperoleh nilai mAP hingga 95,06% pada data *training* serta 50,41% pada data *testing*.

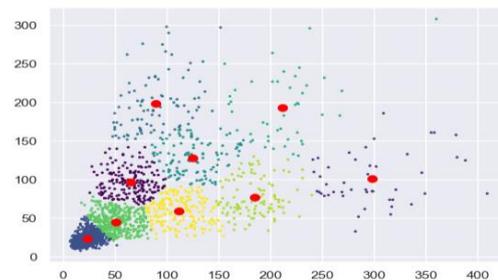
Kata Kunci—Convolutional Neural Network, Deep Learning, K-Means Anchor Box, Mean Average Precision, YOLOv3.

I. PENDAHULUAN

INDONESIA merupakan negara kepulauan terbesar di dunia. Dua pertiga wilayah Indonesia merupakan perairan yang luasnya mencapai 5,8 juta km², terdiri dari Zona Ekonomi Eksklusif (ZEE) 2,7 juta km² dan wilayah laut teritorial 3,1 juta km². Luas wilayah perairan Indonesia tersebut telah diakui sebagai Wawasan Nusantara oleh *United Nation Convention of the Law of the Sea (UNCLOS)* pada tahun 1982 [1]. Bentangan garis pantai Indonesia dengan panjang 81.000 km², menjadikan laut Indonesia dan wilayah pesisir Indonesia memiliki kandungan kekayaan dan sumber daya alam hayati laut yang sangat berlimpah, seperti ikan, terumbu karang, mutiara, mangrove, dan sebagainya. Namun, kondisi ini memicu maraknya kasus pencurian ikan atau dengan kata lain adanya tindak kejahatan yang terjadi di perairan Indonesia yaitu *illegal fishing* [2]. Selain itu, kawasan laut Indonesia juga masih rawan terhadap peristiwa seperti *illegal mining*, *illegal logging*, *drugs trafficking* dan *people smuggling*. Presiden Jokowi telah menetapkan visi kemaritiman yaitu menjadikan negara Indonesia sebagai Poros Maritim Dunia. Mengingat suatu negara dapat dikatakan sebagai Poros Maritim Dunia apabila salah satunya adalah dapat menjaga sumber daya laut dan kedaulatan negaranya. Pengelolaan wilayah perairan Indonesia, terutama di daerah perairan yaitu perbatasan yang dikelola dengan baik, sehingga dapat digunakan untuk kesejahteraan masyarakat Indonesia. Berdasarkan situasi yang terjadi, maka diperlukan suatu sistem baru yang dapat menutupi kelemahan ini yang dapat dilakukan dengan deteksi kapal menggunakan citra digital atau citra satelit.

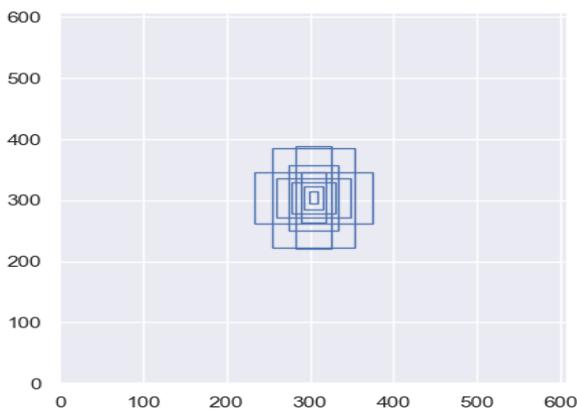


Gambar 1. Performa YOLOv3.

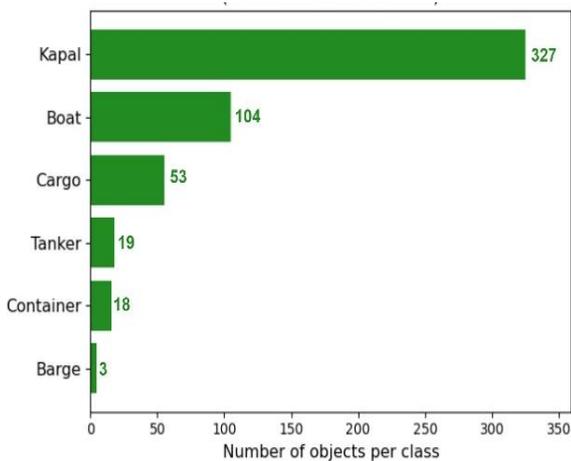


Gambar 2. Plot K-means Hyperparameter Anchor Box.

Convolutional Neural Network (CNN) merupakan salah satu metode *deep learning* yang terus mengalami perkembangan. Penelitian sebelumnya mengenai penerapan CNN pada citra telah dilakukan oleh Alex Krihevsky pada tahun 2012 yang terbukti berhasil mengungguli metode *machine learning* lainnya yaitu *Support Vector Machine (SVM)* [3]. Algoritma *You Only Look Once (YOLO)* adalah sebuah algoritma yang dikembangkan untuk mendeteksi sebuah objek secara *realtime* dengan akurasi yang cukup baik. Sistem pendeteksian yang dilakukan adalah dengan menggunakan *repurpose classifier* atau *localizer* untuk melakukan deteksi. Sebuah model diterapkan pada sebuah citra di beberapa lokasi dan skala. Daerah dengan citra yang diberi *score* paling tinggi akan dianggap sebagai sebuah pendeteksian [4]. Sumber data yang digunakan pada penelitian ini adalah data *remote sensing* pada citra satelit berukuran 768×768 *pixel* sebagai data penelitian yang diunduh dari *Website OneAtlas Sandbox "Ships Detection Machine Learning Dataset"*. Penelitian sebelumnya dengan sumber data yang sama, digunakan data berjumlah 10.000 citra yang diambil secara random, digunakan 2 kategori yaitu kapal dan non kapal dengan menggunakan beberapa metode *machine learning (Naive Bayes, Linear Discriminant Analysis, K-Nearest Neighbors, Random Forest, Support Vector Machine)* dan *deep learning (Convolutional Neural Networks)*, pada metode *machine learning* diperoleh hasil



Gambar 3. Output K-means Hyperparameter Anchor Box.



Gambar 4. Jumlah Objek setiap Kelas.

terbaik dengan metode *Random Forest* yang memiliki hasil akurasi sebesar 93%, sedangkan untuk metode *deep learning* dengan menggunakan *Convolutional Neural Networks* diperoleh akurasi sebesar 94%.

Pada penelitian ini penulis menggunakan metode YOLOv3 sebagai metode untuk ekstraksi fitur pada citra. Penulis mengajukan penggunaan *deep learning* dengan arsitektur YOLOv3 dalam mendeteksi posisi suatu kapal pada citra satelit sehingga dapat dilakukan pendeteksian kapal secara *realtime*. Tujuan yang ingin dicapai dalam penelitian ini adalah memperoleh model arsitektur terbaik YOLOv3 dalam melakukan deteksi kapal dan tipe kapal serta mendapatkan performa pendeteksian keberadaan kapal pada citra satelit dengan teknik *deep learning* terutama dengan arsitektur YOLOv3, dengan memperoleh nilai *loss* dan nilai mAP dari data *training* serta hanya nilai mAP dari data *testing* pada deteksi kapal. Diharapkan cara ini mampu menutupi kelemahan dari sistem VMS yang telah ada dan dapat membantu petugas Angkatan Laut untuk meningkatkan keamanan di perbatasan perairan Indonesia tanpa harus selalu melakukan pengawasan wilayah perbatasan tersebut.

II. TINJAUAN PUSTAKA

A. Citra Digital

Citra digital adalah data digital pada komputer yang merepresentasikan sebuah citra. Sebuah citra yang diwakili oleh $f(a, b)$ berbentuk matriks yang terdiri dari A baris dan B kolom, dimana perpotongan antara kolom dan baris inilah yang disebut *pixel*. Citra digital juga dapat digambarkan

Tabel 1.
Variabel Penelitian

Variabel	Keterangan
$X_{1,ij}$	Nilai Red setiap <i>pixel</i>
$X_{2,ij}$	Nilai Green setiap <i>pixel</i>
$X_{3,ij}$	Nilai Blue setiap <i>pixel</i>
Y_j	Matriks <i>Ground Truth BoundingBox</i>

Tabel 2.
Tipe Kapal

Kelas	Tipe Kapal
0	Boat
1	Cargo
2	Container
3	Tanker
4	Barge
5	Kapal

sebagai fungsi $f(a, b)$ dengan a dan b merupakan koordinat pada sebuah bidang datar yang merepresentasikan kumpulan *pixel* dalam dua dimensi [5].

B. Image Labelling

Image Labelling atau pelabelan gambar atau anotasi gambar adalah proses mengidentifikasi atau mengenali unit yang berbeda dalam suatu gambar [5]. *Bounding box* adalah jenis anotasi pelabelan data yang paling umum digunakan dalam *computer vision* berupa kotak persegi panjang yang digunakan untuk menentukan lokasi objek target. *Bounding box* biasanya diwakili oleh dua koordinat (x_1, y_1) dan (x_2, y_2) atau dengan satu koordinat (x_1, y_1) dan lebar (w) dan tinggi (h) dari kotak pembatas.

C. Convolutional Neural Network (CNN)

CNN didesain untuk mengenali pola visual dari suatu gambar secara langsung dengan proses minimal. CNN pada dasarnya adalah *neural network* dengan banyak *layer*.

D. Transfer Learning

Transfer learning adalah proses menyalin pengetahuan dari jaringan yang sudah ada ke jaringan yang baru untuk menyelesaikan masalah serupa [6]. Teknik ini diimplementasikan untuk mempercepat proses *training* dan meningkatkan performa model [7].

E. You Only Look Once (YOLO)

YOLO adalah sebuah pendekatan untuk sistem pendeteksian objek, yang ditargetkan untuk pemrosesan secara *realtime* dan meringkaskan pendeteksian objek sebagai masalah regresi tunggal, dimana dari *pixel* gambar langsung ke kotak pembatas (*bounding box*) spasial yang terpisah dan probabilitas kelas yang terkait [4].

Berdasarkan Gambar 1 YOLOv3 memiliki performa yang cukup baik dibandingkan dengan metode deteksi objek lainnya dengan waktu pembelajaran yang tercepat.

1) Arsitektur YOLOv3

YOLOv3 menggunakan arsitektur dari Darknet 53 yang artinya memiliki 53 *convlutional layers*. YOLOv3 memiliki 53 lapisan konvolusional yang disebut Darknet 53 yang terdiri dari struktur *convolutional* dan *residual* [4].

2) Anchor Box

Anchor box adalah satu set kotak pembatas yang telah ditentukan dengan tinggi dan lebar tertentu yang didefinisikan untuk menangkap skala dan rasio aspek kelas

Tabel 3.
Hyperparameter Setiap Model

Model	Hyperparameter	Nilai
Kedua model convolutional network	Image Size	416 x 416
	Batch Size	64
	Subdivisions	16
	Training Step	12000
	Learning Rate	0,001
Model1	Anchors	(10x13); (16x30); (33,23); (30,61); (62,45); (59,119); (116,90); (156,198); (373,326)
Model2	Anchors	(11x20); (24x38); (31x50); (42x65); (53x82); (60,85); (88x108); (100x163); (141x168)

objek yang ingin dilakukan deteksi dan biasanya dipilih berdasarkan ukuran objek dalam kumpulan data pelatihan. Dalam deteksi objek perlu dilakukan *k-means clustering* agar dapat disesuaikan *anchor box* yang digunakan pada model dengan *ground truth bounding box* pada *dataset* sehingga memiliki hasil prediksi yang baik.

3) *Intersection over Union (IoU)*

Intersection over union menghitung luas area yang berpotongan lalu membaginya dengan luas area gabungan antar 2 *bounding box*. Nilai IoU digunakan untuk menentukan kesesuaian *bounding box* yang diprediksi dengan luas objek yang sesungguhnya pada citra (*ground truth*). IoU dapat dihitung dengan persamaan sebagai berikut.

$$IoU = \frac{\text{Area Irisan}}{\text{Area Gabungan}} \tag{1}$$

4) *Non Max Supression*

Non Max Supression berhubungan dengan algoritma objek deteksi untuk memfilter *bounding box*. *Non Max Supression* membuat seleksi berdasarkan IoU dengan mengurangi *bounding box* yang muncul secara berlebihan. Seleksi dilakukan hingga tersisa satu saja *bounding box* dan memiliki nilai *confidence* tertinggi.

5) *Fungsi Aktivasi*

Fungsi aktivasi merupakan fungsi yang digunakan pada jaringan saraf untuk mengaktifkan atau tidak mengaktifkan neuron. Hasil perhitungan antara *input*, *weight* dan bias akan dihitung lagi dengan persamaan dari fungsi aktivasi untuk mendapatkan *output* dari setiap layer. Pada penelitian ini, penulis menggunakan fungsi aktivasi ReLU pada *convolutional layer* dan fungsi aktivasi *softmax* pada *output layer* untuk mendapatkan hasil yang merupakan data kategoris. Fungsi ReLU akan merubah nilai *pixel* pada *input* namun tidak merubah dimensi *input* sehingga dimensi *output* berukuran sama dengan *input* [7]. ReLU (*Rectified Linear Unit*) merupakan fungsi aktivasi sebagai solusi dari menghilangnya gradien/*vanishing gradient* dengan cara menerapkan fungsi aktivasi tersebut sebagai persamaan

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

Softmax layer digunakan apabila pada permasalahan *multiclass classification*, *output layer* biasanya memiliki lebih dari satu neuron. Berikut adalah rumus dari fungsi aktivasi softmax.

$$\sigma(a_j) = \frac{\exp(a_j)}{\sum_{k=1}^m \exp(a_k)} \tag{3}$$

dimana $\sum_{j=1}^m \sigma(a_j) = 1$ untuk *softmax*.

6) *Batch Normalization*

Batch Normalization adalah teknik untuk melatih *deep neural networks* dengan melakukan *standardizes input layer* untuk setiap *minibatch*. *Batch normalization* digunakan untuk mempercepat pelatihan jaringan dengan mengurangi *internal covariate shift*. Dimana *internal covariate shift* mengacu pada adanya perubahan distribusi *input* pada *learning system*. *Batch Normalization* berguna untuk meningkatkan kecepatan, performansi dan kestabilan pada jaringan [8]. Berikut persamaan untuk *batch normalization*.

$$\mu_B^{(k)^2} = \frac{1}{m} \sum_{i=1}^m x_i^{(k)} \tag{4}$$

$$\sigma_B^{(k)^2} = \frac{1}{m} \sum_{i=1}^m (x_i^{(k)} - \mu_B^{(k)})^2 \tag{5}$$

$$x_i^{(k)} = \frac{x_i^{(k)} - \mu_B^{(k)}}{\sqrt{\sigma_B^{(k)^2} + \epsilon}} \tag{6}$$

$$y_i^{(k)} = \gamma^{(k)} x_i^{(k)} + \beta^{(k)} = BN_{\gamma^{(k)}, \beta^{(k)}}(x_i^{(k)}) \tag{7}$$

Keterangan:

- $x_i^{(k)}$: input ke-i
- μ_B : mean mini batch
- σ_B^2 : variansi mini batch
- ϵ : numerical stability (konstanta bernilai 1e-5)
- $x_i^{(k)}$: normalize x_i
- $y_i^{(k)}$: output batch normalization

Dimana $k \in [1, d]$ dan $i \in [1, m]$

Berikut merupakan diferensial untuk meminimumkan nilai *loss*.

$$\begin{aligned} \frac{\partial l}{\partial y_i^{(k)}} &= \frac{\partial l}{\partial x_i^{(k)}} \gamma^{(k)} \\ \frac{\partial l}{\partial x_i^{(k)}} &= \sum_{i=1}^m \frac{\partial l}{\partial y_i^{(k)}} x_i^{(k)}, \frac{\partial l}{\partial \beta_i^{(k)}} = \sum_{i=1}^m \frac{\partial l}{\partial y_i^{(k)}} \\ \frac{\partial l}{\partial \sigma_B^{(k)^2}} &= \sum_{i=1}^m \frac{\partial l}{\partial y_i^{(k)}} (x_i^{(k)} - \mu_B^{(k)}) \left(-\frac{\gamma^{(k)}}{2} (\partial \sigma_B^{(k)^2} + \epsilon)^{-3/2} \right) \\ \frac{\partial l}{\partial \mu_B^{(k)}} &= \sum_{i=1}^m \frac{\partial l}{\partial y_i^{(k)}} \frac{-\gamma}{\sqrt{\sigma_B^{(k)^2} + \epsilon}} + \frac{\partial l}{\partial \sigma_B^{(k)^2}} \frac{1}{m} \sum_{i=1}^m (-2)(x_i^{(k)} - \mu_B^{(k)}) \end{aligned} \tag{8}$$

7) *Hyperparameter*

Hyperparameter adalah parameter dari sebuah distribusi di luar distribusi pada model. Dalam konteks model klasifikasi, distribusi model adalah distribusi dari permasalahan klasifikasi yang sedang dipelajari oleh model. Hyperparameter perlu diatur sedemikian rupa dan tepat untuk menghasilkan model dengan performa terbaik [9].

8) Deteksi Objek YOLO

Tabel 4.
Confusion Matrix Model 1 Training

P/G	BO	CA	CO	TA	BA	KA	NO	TO
BO	442	0	0	0	0	1	75	518
CA	0	178	0	0	0	0	0	178
CO	0	0	75	0	0	0	0	75
TA	0	0	0	159	0	0	2	161
BA	0	0	0	0	65	0	1	66
KA	0	3	0	0	0	1136	26	1165
NO	81	14	3	20	7	70		
TO	523	195	78	179	72	1207		

Klasifikasi secara umum adalah proses untuk mengidentifikasi label dari data yang diuji, pada klasifikasi YOLO dilakukan dengan *localization*, yaitu terdapat tambahan pemberian lokasi objek dalam bentuk *bounding box* (bx , by , bh , bw). Langkah-langkah deteksi objek YOLO sebagai berikut.

- Membaca data citra dengan ukuran sebarang.
- Mengubah ukuran citra menjadi 448×448 , lalu membuat *grid* pada citra dengan ukuran $S \times S$ *grids*.
- Memetakan deteksi objek pada setiap *grid cell* dengan *fully connected layer* dan fungsi softmax.
- Melakukan deteksi objek pada masing-masing *grid cell*, masing-masing *grid cell* yang terdapat objek akan bertanggung jawab untuk melakukan deteksi dengan memberikan prediksi *bounding box*.
- Melakukan *non max supression* untuk mendapatkan *bounding box* dengan *confidence score* yang maksimal.

9) Loss Function

Loss adalah nilai dari eror antara nilai hasil prediksi dan nilai sebenarnya. *Loss function* digunakan untuk menghitung nilai dari error. *Loss functions* pada jaringan YOLO didapatkan dari hasil penjumlahan beberapa perhitungan *partial loss functions*. Dengan menghitung kesalahan koordinat, kesalahan klasifikasi, dan kesalahan IoU lalu menjumlahkannya, *total loss function* dapat dinyatakan sebagai berikut,

$$loss = \sum_{i=0}^{s^2} coordErr + clsErr + IoUErr \quad (9)$$

Nilai *loss* untuk prediksi koordinat pusat dinyatakan sebagai berikut,

$$\lambda_{coord} \sum_{i=0}^{s^2} \sum_{j=0}^B 1 [(x_i - x_i\Lambda)^2 + (y_i - y_i\Lambda)^2]_{ij}^{obj} \quad (10)$$

Nilai *loss* untuk prediksi lebar dan tinggi *boundingbox* dinyatakan sebagai berikut,

$$\lambda_{coord} \sum_{i=0}^{s^2} \sum_{j=0}^B 1 \left[(\sqrt{w_i} - \sqrt{w_i\Lambda})^2 + (\sqrt{h_i} - \sqrt{h_i\Lambda})^2 \right] \quad (11)$$

Di mana λ_{coord} menunjukkan faktor pembobot eror pada koordinat dalam *overall loss function* / fungsi keseluruhan *loss*. *Loss* untuk perkiraan kategori dinyatakan sebagai berikut,

$$\sum_{i=0}^{s^2} 1^{obj} \sum_{j=0}^B [(p_i(c) - p_i\Lambda(c))^2] \quad (12)$$

Loss untuk nilai kepercayaan / *confidence* pada prediksi dinyatakan sebagai berikut:

Tabel 5.
Confusion Matrix Model 2 Training

P/G	BO	CA	CO	TA	BA	KA	NO	TO
BO	444	0	0	0	0	1	73	518
CA	0	178	0	0	0	0	0	178
CO	0	0	75	0	0	0	0	75
TA	0	1	0	159	0	0	1	161
BA	0	0	0	0	66	0	0	66
KA	0	3	0	2	0	1132	28	1165
NO	83	23	0	9	8	89		
TO	527	205	75	170	74	1222		

$$\sum_{i=0}^{s^2} \sum_{j=0}^B 1^{obj} [(c_i - c_i\Lambda)^2] + \lambda_{noobj} \sum_{i=0}^{s^2} \sum_{j=0}^B 1^{obj} [(c_i - c_i\Lambda)^2] \quad (13)$$

Dimana C adalah skor kepercayaan/*confidence score*; \hat{C} adalah persimpangan antara *bounding box* yang diprediksi dan data real, ketika sebuah objek ada di dalam sel; maka ℓ_{ij}^{obj} sama dengan 1; jika tidak, itu maka bernilai 0; λ_{noobj} mewakili bo-bot kepercayaan saat tidak ada objek di dalam *bounding box*.

F. Nilai Performa

Nilai evaluasi pertama yang digunakan pada penelitian ini adalah *confusion matrix*. *Confusion matrix* dapat untuk mengetahui hasil kebaikan klasifikasi pada model. Kebaikan klasifikasi dapat dihitung menggunakan nilai *precision* dan *recall* pada masing-masing kelas pada setiap model. Pada permasalahan klasifikasi, menurut kombinasi kelas *ground truth* dan kelas prediktif, hasil dapat dibagi menjadi empat jenis: *true positive* (TP), *false positive* (FP), *true negative* (TN), dan *false negatif* (FN).

Dari hasil *confusion matrix* dapat dihitung *precision* dan *recall* dengan persamaan sebagai berikut.

$$Precision = \frac{TP}{TP+FP} = \frac{TP}{Semua\ Prediksi} \quad (14)$$

$$Recall = \frac{TP}{TP+FN} = \frac{TP}{Semua\ Ground\ Truth} \quad (15)$$

Precision merepresentasikan kemampuan model untuk mengidentifikasi objek terkait sebagai nilai persentase prediksi yang benar. *Recall* adalah kemampuan model untuk menemukan semua objek relevan sebagai nilai persentase positif benar yang dapat terdeteksi di semua *ground truth*. Setelah didapatkan *precision* dan *recall* dapat diperoleh suatu kurva yang disebut sebagai *P-R curve*, dengan *recall* sebagai sumbu horizontal dan *precision* sebagai sumbu vertikal.

Nilai *mean average precision* (mAP) adalah nilai rata-rata dari *average precision* yang merupakan nilai yang didapatkan dari setiap nilai *precision* item relevan yang dihasilkan dan menggunakan nilai 0 untuk item relevan yang tidak dihasilkan oleh sistem. Berikut persamaan untuk menghitung nilai *mean average precision*.

$$mAP = \frac{1}{c} \sum_{i=1}^c AP_i \quad (16)$$

Berikut merupakan persamaan untuk perhitungan AP dengan metode *11-point interpolation*.

$$AP = \frac{1}{11} \sum_{r \in \{0,0.1,\dots,1\}} p_{int\ erp}(r) \quad (17)$$

Tabel 6.
Precision dan Recall Training Model 1 dan Model 2

Kelas	Model 1		Model 2	
	Precision (%)	Recall (%)	Precision (%)	Recall (%)
Boat	84,51	85,33	84,25	85,71
Cargo	91,28	100,00	86,83	100,00
Container	96,15	100,00	100,00	100,00
Tanker	88,83	98,76	93,53	98,76
Barge	90,28	98,48	89,19	100,00
Kapal	94,12	97,51	92,64	97,17
Mean	90,86	96,68	91,07	96,94

$$p_{int\ erp} = \max_{r:\tilde{r}_n} p(\tilde{r}) \tag{18}$$

Perhitungan AP dengan menggunakan metode *all-point interpolation* disajikan pada persamaan berikut.

$$AP = \sum_{n=0} (r_{n+1} - r_n) p_{int\ erp}(r_{n+1}) \tag{19}$$

$$p_{int\ erp}(r_{n+1}) = \max_{r:\tilde{r} \geq r_{n+1}} p(\tilde{r}) \tag{20}$$

Keterangan :

- c* : Jumlah kelas
- R* : item relevan yang dihasilkan oleh sistem
- m* : jumlah item relevan yang dihasilkan suatu kelas
- p* : p resisi
- p_{int erp}* : interpolasi presisi
- r* : recall
- p(ṛ)* : perhitungan presisi pada recall ṛ

G. Kapal

Kapal adalah kendaraan air dengan bentuk jenis apapun, yang digerakkan dengan tenaga mekanik, tenaga angin atau ditunda, termasuk kendaraan yang berdaya apung dinamis, kendaraan di permukaan air, serta alat apung dan bangunan terapung yang tidak berpindah-pindah [10]. Dalam penelitian ini, tipe kapal yang digunakan adalah *boat*, *cargo*, peti kemas (*container*), pengangkut minyak (*tanker*), tongkang/ponton (*barge*).

III. METODOLOGI PENELITIAN

A. Sumber Data

Pada penelitian ini yang digunakan sebagai data penelitian adalah data *testing* yang terdapat pada dataset “*Ships Detection Machine Learning Dataset*” yaitu sebanyak 15.606 file gambar yang diambil pada 20 Januari 2020 pukul 15.20, selanjutnya dari sejumlah file tersebut akan dilakukan anotasi atau pelabelan tipe kapal, sehingga hanya menyisakan 1.465 gambar kapal, sedangkan gambar yang tidak terdapat objek kapal tidak digunakan untuk penelitian.

B. Variabel Penelitian

Variabel penelitian yang digunakan pada penelitian ini terdapat pada Tabel 1 dengan $0 < i \leq K; K$ bernilai 589.824 yang merupakan total *pixel* ($768 \times 768 = 589.824$) dari setiap gambar. $0 < j \leq L; L$ bernilai sejumlah citra satelit yang digunakan.

C. Langkah Analisis

Langkah-langkah analisis pada penelitian ini adalah sebagai berikut.

$$Y_j = \begin{cases} [c_1 \ bx_1 \ by_1 \ bw_1 \ bh_1] & , \text{jika } a = 1 \\ [c_2 \ bx_2 \ by_2 \ bw_2 \ bh_2] & , \text{jika } a = 2 \\ \vdots & \\ [c_n \ bx_n \ by_n \ bw_n \ bh_n] & , \text{jika } a = b \\ [c_b \ bx_b \ by_b \ bw_b \ bh_b] & \end{cases} \tag{21}$$

1) Mengumpulkan data citra kapal.

Gambar terdiri dari panjang dan lebar yang diwakilkan oleh kotak-kotak *pixel* yang tersusun mengikuti ukuran gambar. Setiap *pixel* mewakili satu warna. Warna terbentuk dari kombinasi tiga komponen warna utama yaitu R (*Red*), G (*Green*), dan B (*Blue*). Setiap komponen warna utama memiliki tingkat gradasi yang memiliki nilai antara 0-255, dengan semakin kecil nilai maka warna yang dihasilkan akan semakin gelap dan semakin besar nilai maka warna yang dihasilkan akan semakin terang. Sehingga data yang didapatkan menunjukkan dalam setiap citra terdapat nilai *red* setiap *pixel* ($X_{1,i,j}$), nilai *green* setiap *pixel* ($X_{2,i,j}$), nilai *blue* setiap *pixel* ($X_{3,i,j}$), namun secara realita pada *training* dan *testing* tidak perlu mengubah citra dalam bentuk RGB numerik untuk dapat memproses suatu citra pada YOLOv3, jadi citra dapat diproses secara langsung dalam bentuk gambar.

2) Melakukan pre-processing data.

Menyeleksi gambar untuk data penelitian dengan hanya menggunakan gambar yang terdapat objek kapal dan didapatkan 1.465 gambar yang terdapat objek kapal. Melakukan anotasi pelabelan tipe kapal pada citra gambar dengan 6 tipe kapal yang digunakan yaitu *boat*, *cargo*, *container*, *tanker*, *barge* dan kapal menggunakan bantuan *software labelImg*. Membagi data latih dan uji. Melakukan eksplorasi data dengan terlebih dahulu mengubah format data ground truth voc xml menjadi format csv.

3) Melakukan proses pelatihan deteksi tipe kapal dengan metode YOLO-CNN.

Mendapatkan nilai *anchor box* dengan metode k-means dengan jumlah *cluster anchor* sebanyak 9, dengan langkah-langkah sebagai berikut: (1) Mengubah format *ground truth* yolo txt menjadi format pascal voc xml. Mengubah format *ground truth* pascal voc xml menjadi format coco json. Mendapatkan nilai *anchor box* baru dengan menggunakan *K-means*. Menentukan *hyperparameter* dalam proses pelatihan. Melakukan proses pelatihan. Menghitung nilai *loss* untuk

Tabel 7.
Nilai Rata-Rata Loss dan mAP Training

Model	Avg Loss	mAP (%)
Model 1	0,2388	94,85
Model 2	0,1978	95,06

Tabel 8.
Nilai Average Precision Training Setiap Kelas

Kelas	Average Precision/AP (%)	
	Model 1	Model 2
Boat	80,96	79,60
Cargo	99,94	98,06
Container	98,67	100,00
Tanker	98,69	98,11
Barge	95,02	98,39
Kapal	95,82	96,21

setiap *trainstep* dan mAP untuk setiap model.

4) Memperoleh nilai evaluasi performa untuk kedua model pada data training dan testing sebagai berikut.

Memperoleh *confusion matrix* untuk mengetahui nilai *precision* dan *recall* untuk kedua model. Menghitung nilai mAP pada data untuk kedua model, dengan terlebih dahulu mengubah format data dari *output yolo txt* menjadi partisi beberapa file txt. Mendapatkan grafik mAP dibandingkan dengan *threshold*. Mendapatkan grafik *precision recall curve*. Menghitung model terbaik berdasarkan nilai evaluasi performa pada setiap model.

IV. HASIL DAN PEMBAHASAN

A. Karakteristik Deskripsi Data Hasil Preprocessing

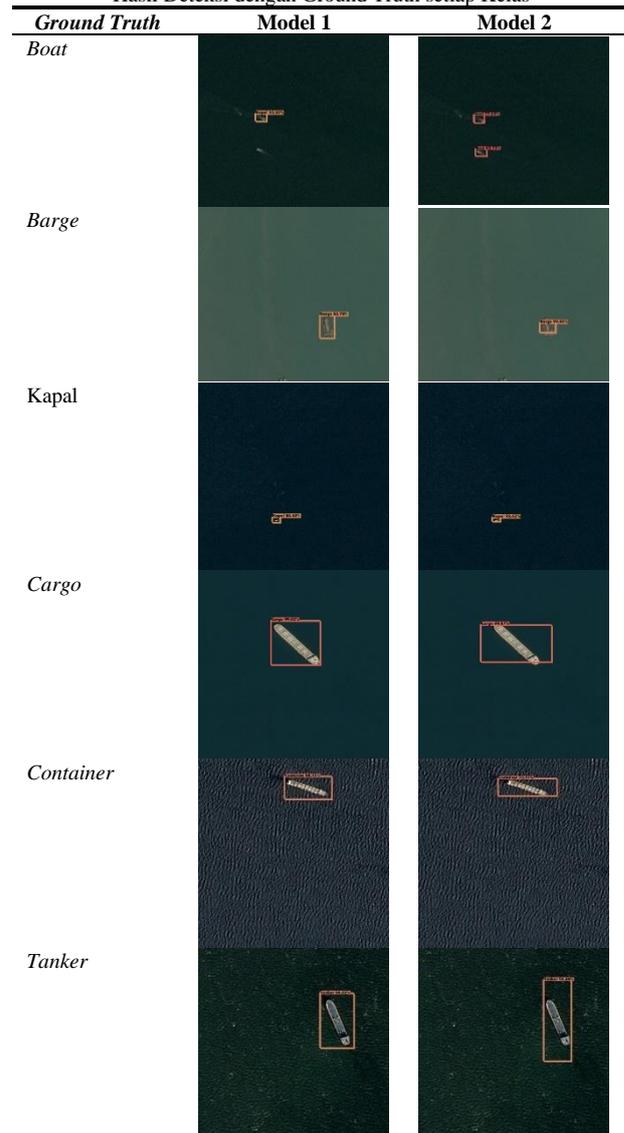
Preprocessing data dalam penelitian ini adalah menyeleksi data citra yang hanya terdapat objek kapal kemudian melakukan anotasi atau pelabelan gambar sebagai data *ground truth bounding box*. Anotasi dilakukan untuk menandai dengan menggunakan *bounding box* pada objek kapal yang terdapat pada citra satelit, sehingga diperoleh data yang terdiri dari tipe kapal beserta lokasi dan spesifikasi *bounding box* pada citra sebagai berikut. Pada Tabel 2 dapat diketahui 6 kelas tipe kapal yaitu sebagai berikut.

Anotasi atau pelabelan tipe kapal dilakukan dengan melihat karakteristik kapal secara visual, setiap tipe kapal memiliki ciri yang berbeda dari warna, bentuk, dan karakteristik lainnya. Dataset yang digunakan dalam penelitian ini yaitu terdiri dari 1465 gambar, 6 kategori kapal, dan 2687 kapal dalam dataset citra satelit. Dilakukan pembagian 2 kelompok data citra satelit hasil dari seleksi data dan *preprocessing* yaitu data *training* dan *testing*. Dari seluruh data kapal pada citra satelit dilakukan pembagian data sebanyak 80% data yaitu 1172 citra satelit sebagai data *training* dan sebanyak 20% data yaitu 293 citra satelit sebagai data *testing*. Data *training* digunakan untuk menangkap pola-pola yang dimiliki oleh citra satelit sehingga diperoleh model terbaik dengan hasil perhitungan nilai mAP dan *loss*. Data *testing* dipergunakan untuk melakukan evaluasi dengan menggunakan data yang baru dengan mendapatkan nilai mAP.

B. Klasifikasi dan Lokalisasi Tipe Kapal YOLO-CNN

Metode klasifikasi dan lokalisasi tipe kapal yang digunakan adalah metode *You Only Look Once (YOLO)*. Pada proses pelatihan dilakukan *transfer learning* arsitektur YOLOv3 yang disebut juga Darknet-53, yang dilakukan pada

Tabel 9.
Hasil Deteksi dengan Ground Truth setiap Kelas



data *training* sehingga diperoleh model *custom* sesuai data yang digunakan. YOLOv3 memprediksi 3 skala klasifikasi dan lokalisasi tipe kapal yang berbeda, jadi jika kita melakukan *input* gambar berukuran 416 x 416, akan menghasilkan 3 *output* yaitu 13 x 13 x 33, 26 x 26 x 33, dan 52 x 52 x 33. Klasifikasi dan lokalisasi tipe kapal dengan 3 skala yang dilakukan oleh arsitektur YOLOv3 memiliki tujuan agar dapat dilakukan pada berbagai ukuran objek, baik objek berukuran besar, sedang, dan kecil dengan lebih baik. Pada setiap skala pada YOLOv3, setiap sel grid memprediksi 3 *bounding box* menggunakan 3 *anchors*, sehingga membuat jumlah total *anchors* yang digunakan sebanyak 9 *anchors*.

Pada proses *training* digunakan dua model jaringan konvolusional yang dipertimbangkan, yaitu YOLOv3 dengan 9 *anchors* default serta YOLOv3 dengan 9 *anchors* yang dilakukan proses *k-means*, keduanya dilakukan pelatihan *transfer learning* pada jaringan arsitektur Darknet-53, Tabel 3 menunjukkan *hyperparameter* yang digunakan untuk dua model dalam penelitian.

Pada penelitian ini dilakukan *training* dengan membandingkan beberapa *hyperparameter* yang digunakan karena perbedaan nilai *hyperparameter* mempengaruhi kecepatan dan kualitas proses pembelajaran yang dapat

Tabel 10.
Nilai Average Precision Training Setiap Kelas

Kelas	Average Precision/AP (%)	
	Model 1	Model 2
Boat	12,01	23,66
Cargo	68,25	70,58
Container	66,83	70,52
Tanker	60,10	79,43
Barge	10,00	22,00
Kapal	34,11	36,27
mAP (%)	41,88	50,41

berpengaruh pada pembentukan model sehingga diperoleh model yang terbaik. Pada proses *training* digunakan *batch size* sebesar 64 dan jumlah maksimal *train step* 12000. *Batch size* dengan nilai 64 menyebabkan terambilnya 64 gambar sebagai sampel dan *update weight* dan bias dalam 1 *train step*. *Subdivisions* atau *minibatch* dengan nilai 16 artinya dalam 1 *subdivisions/minibatch* diambil 4 citra. *Train step/iterasi* yang terdiri dari *update weight* dan bias berdasarkan *batch size* dan pengecekan rata-rata *loss* dan akurasi untuk data *training*. Total banyak pelatihan citra yang dilakukan untuk memperoleh satu model adalah sebanyak total 768000 citra. Perbedaan kedua model terletak pada *hyperparameter anchors* dengan jumlah *anchors* yang sama-sama berjumlah 9, namun memiliki nilai lebar dan tinggi *anchors* yang berbeda. Pada Model 1 nilai *anchors* merupakan nilai *anchors* default yang diperoleh dari pelatihan YOLOv3 yang berjumlah 9 dengan ukuran lebar dan tinggi (10x13); (16x30); (33,23); (30,61); (62,45); (59,119); (116,90); (156,198); dan (373,326). Pada Model 2 nilai *anchors* diperoleh dari perhitungan menggunakan metode *k-means* dengan $k=9$ seperti halnya default pada YOLOv3 dengan hasil sebagai berikut.

Dalam Gambar 2 diperoleh plot pada koordinat x yang merupakan nilai lebar (*bw*) untuk sejumlah *ground truth boundingbox*, sedangkan pada koordinat y merupakan nilai tinggi (*bh*) untuk setiap *ground truth boundingbox*. Selanjutnya dilakukan *k-means* dengan $k=9$ sehingga diperoleh 9 *centroid* sebagai pusat *cluster*. Diperoleh nilai 9 *anchor box* sebagai *output* dari *k-means hyperparameter anchor box* seperti yang ditunjukkan pada Gambar 3.

Pada Gambar 3 diperoleh hasil dari *k-means anchor box* berjumlah 9 *anchor box* baru yang digunakan sebagai *hyperparameter anchors* pada Model 2. Tujuan dari klusterisasi ini yaitu dapat disesuaikan *anchor box* yang digunakan dengan *ground truth boundingbox* pada dataset yang digunakan sehingga memiliki hasil yang baik. Sehingga dalam penelitian untuk mengetahui performa kebaikan pelatihan digunakan model dengan variasi *hyperparameter anchors* yaitu *default anchor box* pada YOLOv3 pada *coco dataset* pada Model 1 dibandingkan dengan model yang dilakukan *k-means anchor box* pada Model 2.

Nilai evaluasi pertama pada penelitian ini adalah *confusion matrix*, yang digunakan untuk mengetahui hasil kebaikan klasifikasi pada kedua model, selanjutnya dapat dihitung nilai *precision* dan *recall* pada masing-masing kelas setiap model. Nilai evaluasi kedua yaitu mAP yang merupakan nilai evaluasi yang paling populer pada deteksi objek, nilai mAP dapat digunakan untuk membandingkan kebaikan model dengan model lainnya pada tugas deteksi objek. Tabel 4 dan Tabel 5 merupakan *output confusion matrix* untuk data

training pada Model 1 dan Model 2 yang disajikan untuk mengetahui hasil kinerja klasifikasi.

Berdasarkan hasil *confusion matrix* Model 1 pada Tabel 6, kelas BO (*Boat*) memiliki *True Positive* sebanyak 442, *False Negative* sebanyak 76, *False Positive* sebanyak 81; kelas CA (*Cargo*) memiliki *True Positive* sebanyak 75, *False Negative* sebanyak 0, *False Positive* sebanyak 17; kelas CO (*Container*) memiliki *True Positive* sebanyak 75, *False Negative* sebanyak 0, *False Positive* sebanyak 3; kelas TA (*Tanker*) memiliki *True Positive* sebanyak 159, *False Negative* sebanyak 2, *False Positive* sebanyak 20; kelas BA (*Barge*) memiliki *True Positive* sebanyak 65, *False Negative* sebanyak 1, *False Positive* sebanyak 65; kelas KA (*Kapal*) memiliki *True Positive* sebanyak 1136, *False Negative* sebanyak 29, *False Positive* sebanyak 71.

Pada Tabel 5, diperoleh hasil *confusion matrix* pada Model 2, kelas *Boat* memiliki *True Positive* sebanyak 444, *False Negative* sebanyak 74, *False Positive* sebanyak 83; kelas *Cargo* memiliki *True Positive* sebanyak 178, *False Negative* sebanyak 0, *False Positive* sebanyak 27; kelas *Container* memiliki *True Positive* sebanyak 75, *False Negative* sebanyak 0, *False Positive* sebanyak 0; kelas *Tanker* memiliki *True Positive* sebanyak 159, *False Negative* sebanyak 2, *False Positive* sebanyak 11; kelas *Barge* memiliki *True Positive* sebanyak 66, *False Negative* sebanyak 0, *False Positive* sebanyak 8; kelas *Kapal* memiliki *True Positive* sebanyak 1132, *False Negative* sebanyak 33, *False Positive* sebanyak 90. Dari hasil *confusion matrix*, dapat dihitung *precision* dan *recall* pada masing-masing kelas sebagai berikut.

Pada Tabel 6, diperoleh hasil perhitungan *precision* dan *recall* untuk masing-masing kelas pada Model 1 dan Model 2. *Precision* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan hasil yang diprediksi positif. Mean *precision* Model 2 lebih tinggi dibandingkan dengan Model 1, artinya secara rata-rata Model 2 lebih baik dalam melakukan prediksi tipe kapal dengan lebih akurat. *Recall* merupakan rasio prediksi benar positif dibandingkan dengan keseluruhan data yang benar positif. Mean *recall* Model 2 lebih tinggi dibandingkan dengan Model 1, artinya secara rata-rata Model 2 lebih baik dalam mengklasifikasikan jenis kapal yang sesuai dengan *ground truth*.

Hasil kinerja klasifikasi yang didapatkan pada metode YOLOv3 dengan menggunakan *Loss* dan mAP disajikan pada Tabel 7.

Tabel 7 menunjukkan kinerja klasifikasi yang didapatkan dari setiap model. Kinerja klasifikasi terbaik terdapat pada model 2 dengan model yang dilakukan *training* menggunakan *anchors* sebanyak 9 *anchors* yang melalui proses *k-means*. Hasil ini mendapatkan rata-rata *loss* model 2 pada data *training* lebih kecil daripada model 1 yaitu senilai 0,1978, artinya nilai bias prediksi terhadap lokalisasi *bounding box*, *confidence score* serta klasifikasi secara keseluruhan pada model 2 yaitu sebesar 0,1978. Nilai mAP pada model 2 pada data *training* lebih tinggi dibandingkan model 1 yaitu senilai 95,06%, nilai tersebut memiliki arti bahwa rata-rata ketepatan prediksi kapal pada setiap kelas yaitu sebesar 95,06%. Penggunaan *anchors* dengan menggunakan metode *k-means* mengindikasikan adanya penambahan kinerja klasifikasi.

Selanjutnya ditampilkan kinerja kedua model pada setiap kelas. Tabel 8 merupakan hasil kinerja untuk setiap kelas dengan metode YOLOv3. Model 1 dapat melakukan prediksi tipe kapal yang lebih tinggi untuk tipe kapal “Boat” dan “Cargo”, sedangkan model 2 dapat melakukan prediksi tipe kapal yang lebih tinggi untuk tipe kapal “Container”, “Tanker”, “Barge”, dan “Kapal”. Pada Tabel 5 didapatkan bahwa model terbaik yaitu model 2. Sehingga berdasarkan beberapa nilai evaluasi performa yang diperoleh dapat disimpulkan bahwa model terbaik dengan metode YOLOv3 adalah Model 2 dengan menggunakan *hyperparameter anchors* sebanyak 9 yang dilakukan *k-means clustering*.

C. Deteksi Tipe Kapal

Deteksi tipe kapal untuk kedua model dengan *hyperparameter* yang didapat digunakan sebagai acuan untuk melakukan deteksi citra satelit dengan objek kapal pada data *testing*. Berikut ditampilkan *ground truth bounding box* untuk setiap kelas pada data *testing*.

Berdasarkan Gambar 4 kelas “Kapal” memiliki jumlah objek yang tertinggi pada data *testing* yaitu sebanyak 327 objek, kelas “Boat” sebanyak 104 objek, kelas “Cargo” sebanyak 53 objek, kelas “Tanker” sebanyak 19 objek, kelas “Container” sebanyak 18 objek, dan kelas “Barge” sebanyak 3 objek. Kedua model dibandingkan hasil deteksi pada setiap kelas.

Dari Tabel 9, ditunjukkan bahwa pada masing-masing data gambar memiliki hasil deteksi yang berbeda-beda dari segi spesifikasi ukuran dan koordinat *boundingbox*, serta *confidence score* pada kelas yang diprediksi. *Confidence score* yang diperoleh hampir pada semua gambar pada setiap kelas yang ditampilkan cenderung pada model 2 didapatkan *confidence score* yang lebih tinggi dibandingkan dengan pada model 1. Pada kelas dengan *ground truth bounding box* “Boat” pada model 1 terdapat kesalahan deteksi, dimana pada *ground truth bounding box* yang seharusnya merupakan “Boat” namun dideteksi sebagai “Kapal”. Hal ini disebabkan karena adanya karakteristik *pixel* yang hampir mirip antara kelas “Boat” dengan kelas “Kapal”. Untuk mengetahui performa klasifikasi digunakan nilai mAP. Sebelum menghitung nilai mAP dibentuk *confusion matrix* yang menunjukkan kemampuan model dalam mengklasifikasikan tipe kapal. Pada data *testing confusion matrix* tidak ditampilkan karena memiliki perhitungan yang sama dengan *confusion matrix* pada data *training*. Oleh karena itu berikut merupakan daftar *average precision* untuk setiap kelas.

Berdasarkan Tabel 10 diperoleh perbandingan *average precision* untuk kedua model, dimana model 2 tetap merupakan model yang terbaik, dengan hasil *average precision*

pada setiap kelas memiliki nilai yang lebih tinggi jika dibandingkan dengan model 1, sehingga diperoleh nilai mAP pada model 2 juga lebih tinggi jika dibandingkan dengan model 1, yaitu masing-masing sebesar 50,41% dan 41,88%. Oleh karena itu, model 2 model yang menunjukkan kinerja terbaik dari kedua model.

V. KESIMPULAN

Berdasarkan analisis dan pembahasan yang telah dilakukan, kesimpulan yang dapat ditarik adalah metode yang memiliki kinerja klasifikasi yang terbaik untuk melakukan deteksi tipe kapal pada citra satelit merupakan model 2 dengan model yang dilakukan *training* menggunakan *anchors* sebanyak 9 *anchors* dengan melalui proses *k-means*. Hasil ini mendapatkan rata-rata *loss* pada model 2 lebih kecil daripada model 1 yaitu senilai 0,1978, serta nilai mAP pada model 2 pada data *training* lebih tinggi dibandingkan model 1 yaitu masing-masing senilai 95,06% dan 94,85%. Hasil deteksi tipe kapal pada data *testing* diperoleh hasil *average precision* pada model 2 setiap kelas memiliki nilai yang lebih tinggi jika dibandingkan dengan model 1, sehingga diperoleh nilai mAP pada model 2 lebih tinggi dibandingkan model 1, yaitu masing-masing sebesar 50,41% dan 41,88%.

DAFTAR PUSTAKA

- [1] A. Q. Jaelani and U. Basuki, “Illegal unreported and unregulated (IUU) fishing: upaya mencegah dan memberantas illegal fishing dalam membangun poros maritim indonesia,” *J. Kaji. Ilmu Huk.*, vol. 3, no. 1, 2014.
- [2] Y. D. Pratiwi, “Pertanggung jawaban pidana illegal fishing korporasi dalam cita-cita indonesia poros maritim dunia,” *J. Defendonesia*, vol. 1, no. 2, p. 66, 2016.
- [3] W. S. Eka Putra, “Klasifikasi citra menggunakan convolutional neural network (CNN) pada caltech 101,” *J. Tek. ITS*, vol. 5, no. 1, 2016, doi: 10.12962/j23373539.v5i1.15696.
- [4] R. Huang, J. Pedoeem, and C. Chen, “Yolo-Lite: a Real-Time Object Detection Algorithm Optimized for Non-GPU Computers,” in *IEEE International Conference on Big Data*, 2018, pp. 2503–2510, doi: 10.1109/BigData.2018.8621865.
- [5] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3th ed. New Jersey: Pearson Education, 2008.
- [6] M. Sewak, R. Karim, Md, and P. Pujari, *Practical Convolutional Neural Networks*. Birmingham: Packt Publishing, 2018.
- [7] J. Petterson, *Deep Learning: A Practitioner’s Approach*, 1st ed. Massachusetts: O’Reilly Medai, 2017.
- [8] S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” in *Proceedings of the 32nd International Conference on Machine Learning*, 2015, vol. 37, pp. 448–456.
- [9] J. Bergstra, B. Kégl, Y. Bengio, and R. Bardenet, “Alogarithm for Hyper-Parameter Optimizaion,” in *Advances in Neural Information Processing System*, 2011.
- [10] Kementerian Luar Negeri, *Undang-Undang Republik Indonesia No. 17 Tahun 2008*. Jakarta: Kementerian Luar Negeri, 2008.