

**Deteksi Kondisi Gigi Manusia pada Citra Intraoral Menggunakan YOLOv5****Ahmad Fauzi Makarim<sup>1</sup>, Tita Karlita<sup>1</sup>, Riyanto Sigit<sup>1</sup>, Bima Sena Bayu Dewantara<sup>1</sup>, Arya Brahmanta<sup>2</sup>**

fauzimakarim@pasca.student.pens.ac.id, tita@pens.ac.id, riyanto@pens.ac.id, bima@pens.ac.id, arya.brahmanta@hangtuah.ac.id

<sup>1</sup>Departemen Teknik Informatika dan Komputer, Politeknik Elektronika Negeri Surabaya<sup>2</sup>Departemen Orthodontic, Universitas Hang Tuah**Informasi Artikel**

Diterima : 17 Ags 2023

Direview : 19 Ags 2023

Disetujui : 29 Ags 2023

**Kata Kunci**Deteksi Objek, YOLOv5,  
Intraoral, Gigi**Abstrak**

Proses identifikasi dan pencatatan rekam medis pada praktik kedokteran gigi masih dilakukan secara manual. Akibatnya, proses tersebut memakan waktu yang cukup lama. Pada penelitian ini metode deteksi objek dimanfaatkan untuk membantu dokter melakukan identifikasi pada gigi pasien. YOLOv5 dipilih untuk dilatihkan pada dataset citra intraoral dengan lima kelas kondisi gigi (normal, karies, tumpatan, sisa akar, dan impaksi). Dataset yang digunakan berjumlah 1.767 data citra intraoral yang diambil dan dilabeli oleh dokter gigi. Dataset dibagi menjadi tiga bagian, 10% digunakan untuk data *testing* dan 90% digunakan untuk data *training* dan *validation*. Dilakukan komparasi performa berdasarkan nilai metrik evaluasi terhadap tiga jenis model YOLOv5 (S, M, L). Dari hasil pelatihan, YOLOv5 M sebagai model terbaik mendapatkan nilai mAP sebesar 84%, dan 82% nilai akurasi *testing*. Penelitian ini telah memenuhi tujuan utama untuk membangun sebuah model *deep learning* yang robust untuk mendekripsi dan mengklasifikasi beberapa kondisi gigi pada manusia.

**Keywords**Object Detection, YOLOv5,  
Intraoral, Teeth**Abstrak**

*The process of identifying and recording medical records in dental practice is still done manually. As a result, the process takes quite a long time. In this study, the object detection method was used to help doctors identify the patient's teeth. YOLOv5 was selected for training on an intraoral image dataset with five classes of tooth condition (normal, caries, restoration, gangren radix, and impaction). The dataset used is 1,767 intraoral images taken and labeled by dentists. The dataset is divided into three parts, 10% is used for data testing and 90% is used for data training and validation. Performance comparisons were made based on evaluation metric values for the three types of YOLOv5 models (S, M, L). From the results of the training, YOLOv5 M as the best model obtained a mAP value of 84% and 82% testing accuracy. This research has fulfilled the primary objective of building a robust deep learning model to detect and classify several dental conditions in humans.*

## A. Pendahuluan

Seorang dokter berkewajiban untuk mengisi rekam medis setiap gigi pasiennya. Pada rekam medis kedokteran gigi, kondisi gigi manusia dibagi menjadi 26 jenis. Pada praktik kedokteran gigi, pemeriksaan dan perekaman rekam medis masih dilakukan secara manual. Akibatnya, banyak kasus kurangnya kelengkapan pengisian rekam medis karena waktu yang dibutuhkan cukup lama untuk mengevaluasi dan mengisi rekam medis [1][2].

Dalam beberapa tahun terakhir, *artificial intelligence*, khususnya *deep learning*, telah banyak diimplementasikan untuk membantu dokter gigi dalam menganalisis gigi pasiennya. Namun, penelitian mendalam mengenai penggunaan *deep learning* dalam kedokteran gigi masih terbatas, khususnya untuk mengenali kondisi gigi menggunakan citra intraoral [3][4]. Gambar intraoral dapat diambil langsung menggunakan kamera intraoral. Berbeda dengan citra radiografi yang harus diproses terlebih dahulu.

CNN (*Convolutional Neural Network*) merupakan salah satu metode dalam *deep learning* yang biasa digunakan untuk menyelesaikan masalah klasifikasi citra dalam kegiatan medis [5]. Kebanyakan pendekatan CNN digunakan untuk deteksi, klasifikasi, dan segmentasi objek. Luya Lian et. al. [6] menggunakan metode *deep learning* dengan model nnU-Net dan DenseNet121 untuk mendekripsi dan mengklasifikasi kondisi gigi karies. Kedua model tersebut dilakukan *training* menggunakan dataset gambar gigi berbentuk *panoramic film*. Terdapat total 1071 data yang digunakan dalam proses *training* dan *validation*. Untuk mengetahui performa model dalam menangani kasus nyata, model diujicoba oleh enam dokter gigi. Dari 89 data test yang diambil diluar dari data *training*, kedua model yang digunakan mendapatkan hasil yang sangat memuaskan dimana hampir tidak ada perbedaan antara hasil deteksi dan klasifikasi antara menggunakan metode *deep learning* dengan hasil pemeriksaan keenam dokter gigi. Model nnU-Net mendapatkan akurasi sebesar 98% dan model DenseNet121 mendapatkan akurasi sebesar 95%.

Moutselos et. al. [7] secara spesifik menggunakan pendekatan CNN untuk tujuan segmentasi dalam mengenali gigi karies pada manusia. Penelitiannya bertujuan untuk medeteksi serta mengklasifikasi gigi karies pada permukaan gigi yang terhalang berdasarkan skala 7 kelas ICDAS (International Caries Detection and Assessment System). Dataset yang digunakan pada penelitiannya adalah gambar yang diambil dengan menggunakan kamera intraoral dan berjumlah 88 data. Dataset dilatihkan menggunakan model Mask R-CNN dan dilakukan beberapa augmentasi. Penelitiannya menghasilkan akurasi sebesar 88%.

Pada tahun 2021 Meyer et. al. [8] melakukan penelitian untuk mendekripsi gigi karies menggunakan metode *deep learning*. CNN dipilih untuk mendekripsi gigi karies dengan dilatihkan menggunakan 2.417 dataset yang berasal dari gambar kamera intraoral dengan 1.317 gambar dengan *occlusal* dan 1.100 gambar tanpa *occlusal*. MobileNetV2 dipilih menjadi *base model* pada *network* yang digunakan dalam proses *training*. Dari hasil proses *training* dihasilkan bahwa model dapat memberikan akurasi 93% dalam mengklasifikasi gigi karies. Hasil tersebut juga dibandingkan dengan hasil pemeriksaan ekspert tenaga medis gigi.

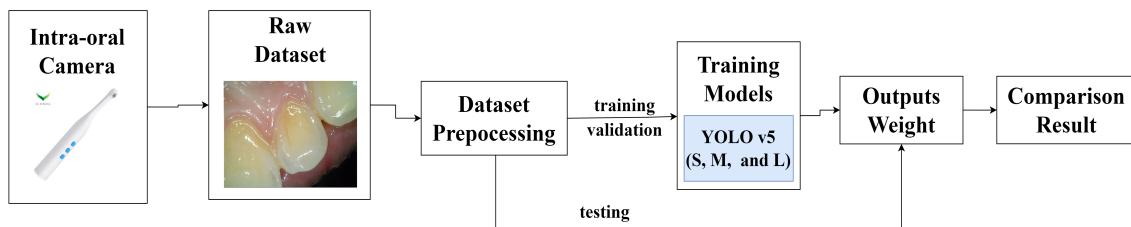
Penelitian lain menggunakan metode *deep learning* untuk mengklasifikasi gigi karies juga dirujuk oleh Xuang Zhang et. al. pada tahun 2020 di China [9]. CNN

digunakan untuk *training* kepada 3.932 gambar dari kamera intra oral yang berasal dari 625 relawan. Model CNN dibangun dengan mengadaptasi *Single Shot MultiBox Detector*. Dari hasil penelitian ini didapatkan bahwa model yang dibangun berhasil mendapatkan nilai *area under curve* (AUC) sebesar 85% dengan nilai *confidence* sebesar 95% dalam mengklasifikasi gigi berlubang.

Strategi digitalisasi untuk kesehatan gigi dan mulut dapat diimplementasi untuk membantu dokter menganalisa gigi pasien[10]. Analisa manual dalam penanganan masalah gigi memerlukan waktu dan kompetensi. Resiko kesalahan prediksi akibat kelelahan dan keteledoran manusia dapat terjadi pada analisa manual. Pendekatan komputerisasi untuk mengidentifikasi masalah kesehatan gigi dapat dijadikan diagnosa dini yang dapat menghetikan beberapa masalah lebih serius[11].

Pada paper ini kami mengusulkan solusi untuk membantu dokter gigi megidentifikasi kondisi gigi pasiennya menggunakan *deep learning*. Tujuan utama dalam penelitian ini adalah untuk membangun dataset citra intraoral dan membuat model yang robust dalam mendeteksi lima kondisi gigi manusia, yaitu : normal, karies, tumpatan, sisa akar, dan impaksi. Dataset khusus dibangun dengan bantuan para dokter gigi untuk mengumpulkan dan melabeli gambar intraoral. Total dataset yang dikumpulkan sebanyak 1.767 gambar intraoral. Metode augmentasi dengan *horizontal flip* digunakan karena tidak mengurangi validitas keadaan gigi secara asli untuk memperkaya variasi dataset. Kemudian tiga jenis model YOLOv5 (S, M, L) dilatihkan untuk mendeteksi dan mengenali kelima jenis kondisi gigi tersebut.

## B. Metode Penelitian



**Gambar 1.** *Design system Penelitian*

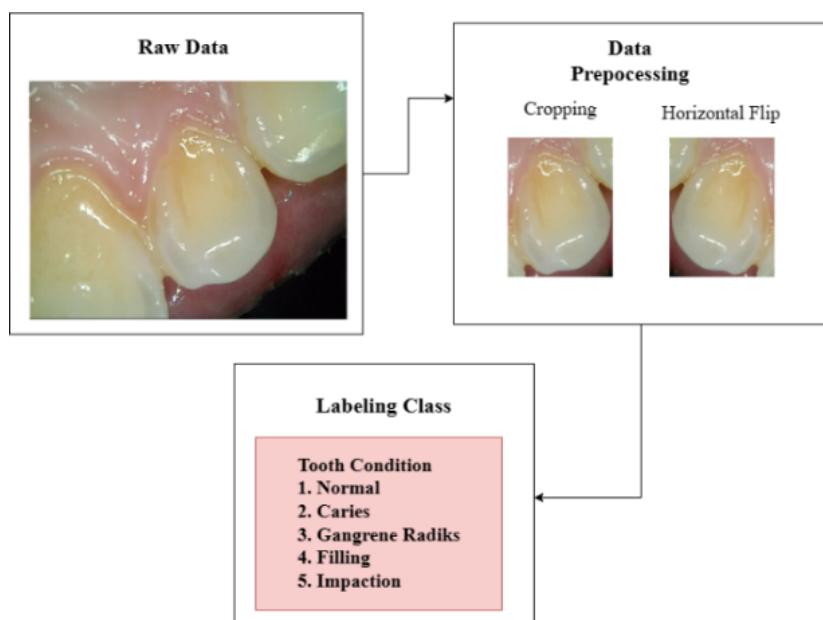
Gambar 1 merupakan *design system* untuk membangun model yang dapat mendeteksi lima jenis kondisi gigi. Dataset diambil dengan menggunakan kamera intraoral oleh ahli gigi. Dataset mentah kemudian diolah dengan beberapa *preprocessing* data sebelum dilatihkan kepada tiga jenis model YOLOv5. *Weight* yang dihasilkan dari proses *training* ketiga model dibandingkan performanya dengan beberapa metrik evaluasi.

### Dataset

Tahapan *preprocessing* dataset terdiri dari pengumpulan, manajemen, anotasi, dan augmentasi. Dataset diambil secara langsung menggunakan kamera intraoral oleh dokter gigi. Citra intraoral mentah dari hasil tangkapan intraoral kamera diorganisir sesuai dengan kelasnya masing – masing. Kondisi detail setiap kelas kondisi gigi ditunjukkan pada Gambar 2.

**Gambar 2.** Kelas Kondis Gigi

Anotasi dataset dilakukan dengan bantuan tenaga medis kedokteran gigi dengan menggunakan *Labelimg tools*. Augmentasi juga dilakukan untuk memperkaya variasi dataset yang ada. *Horizontal flip* digunakan sebagai metode augmentasi karena tidak menghilangkan validitas gigi. Gigi manusia memiliki bentuk mahkota yang sama secara horizontal dan berbeda secara vertical oleh karena itu *horizontal flip* dipilih sebagai metode augmentasi [12]. Sebelumnya dataset dipotong secara manual sehingga dalam satu gambar hanya ada satu objek gigi. Proses penyiapan data ditunjukkan pada Gambar 3.

**Gambar 3.** Preprocessing Dataset

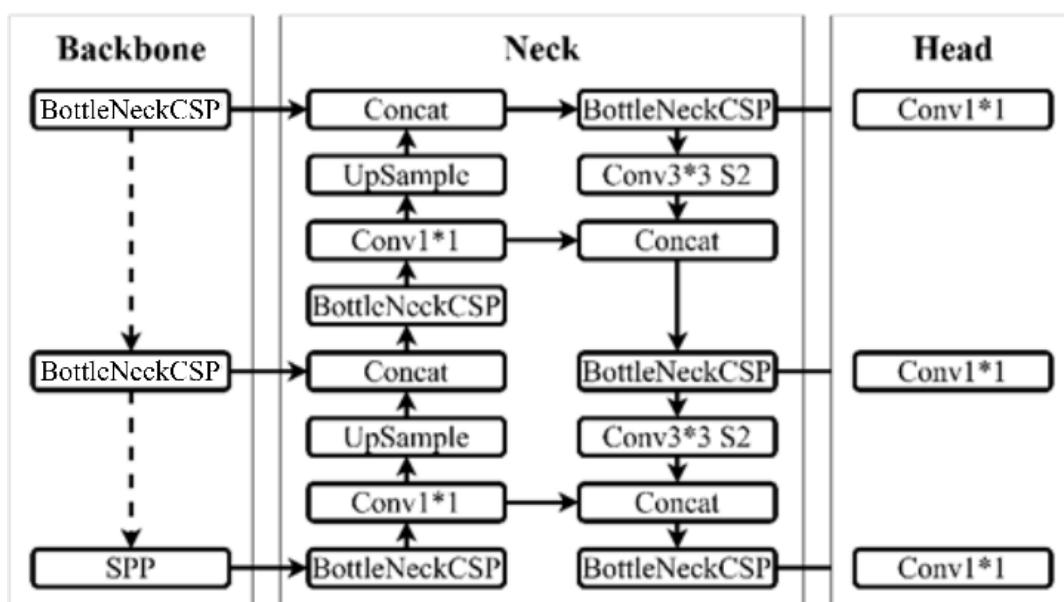
Total dataset hasil augmentasi ditunjukkan pada tabel 1. Dataset dibagi menjadi tiga bagian yaitu : data *training*, data *validation*, dan data *testing*. Data *testing* berjumlah 10% dari seluruh dataset. Sedangkan data *training* dan *validation* berjumlah 80% dan 20% dari 90% total dataset.

**Tabel 1.** Total Dataset

Class	Total
Impaksi	122
Karies	500
Normal	575
Sisa Akar	115
Tumpatan	455

## Arsitektur YOLOv5

YOLO v5 dikembangkan oleh Glenn Jocher sebagai pengembangan lebih lanjut dari versi sebelumnya. Tujuan utama YOLOv5 adalah untuk mempercepat dan menyempurnakan performa deteksi objek dengan mengoptimalkan arsitektur dan teknik *training*. Terdapat 3 bagian penting pada arsitektur YOLOv5 yaitu : *backbone*, *neck*, dan *head*. *Backbone* merupakan bagian dari arsitektur yang berfungsi sebagai pengekstrasi fitur. Contoh model dari *backbone* seperti CSPDarknet53, VGG16, SqueezeNet, MobileNet, dan ShuffleNet. Pada penelitian ini arsitektur *backbone* yang digunakan pada YOLOv5 adalah CSPDarknet53 [13]. *Neck* adalah bagian *pipeline* yang mengagregasi *feature maps* dari bagian *backbone*. *Head* memiliki tanggung jawab dalam melakukan prediksi dari *feature maps* [14]–[16]. Secara lebih spesifik arsitektur YOLOv5 yang digunakan pada paper ini ditunjukkan pada Gambar 4.



**Gambar 4.** Arsitektur YOLOv5

Pada *design system* ditunjukkan bahwa terdapat tiga jenis YOLOv5 yang dilatihkan pada dataset yaitu YOLOv5 S, M, dan L. Perbedaan tiga jenis model tersebut adalah pada parameter kedalaman model dan *channel*. YOLOv5 S memiliki nilai kompleksitas arsitektur yang paling rendah, sehingga model lebih ringan saat diimplementasikan. Akan tetapi, biaya yang harus dibayar adalah performa deteksi akan menurun dan bersifat sebaliknya dengan model lainnya.

## Evaluation Matrices

Beberapa matriks evaluasi digunakan untuk mengukur performa hasil proses *training* dari ketiga jenis model. *Precision*, *recall*, *f1 score*, dan *mAP* (*mean average precision*) merupakan matrik evaluasi yang digunakan pada penelitian ini untuk mengukur performa dari model. Terdapat beberapa komponen dalam pengukuran matrik evaluasi, sebagai berikut :

- *True Negative* (TP) : Model memprediksi objek bernilai positif dan pada kenyataannya objek benar bernilai positif.

- *False Negative (FN)* : Model memprediksi objek bernilai negatif dan pada kenyataannya objek bernilai positif.
- *False Positive (FP)* : Model memprediksi objek bernilai positif dan pada kenyataannya objek bernilai negatif.

$$\text{precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{recall} = \frac{TP}{TP + FN} \quad (2)$$

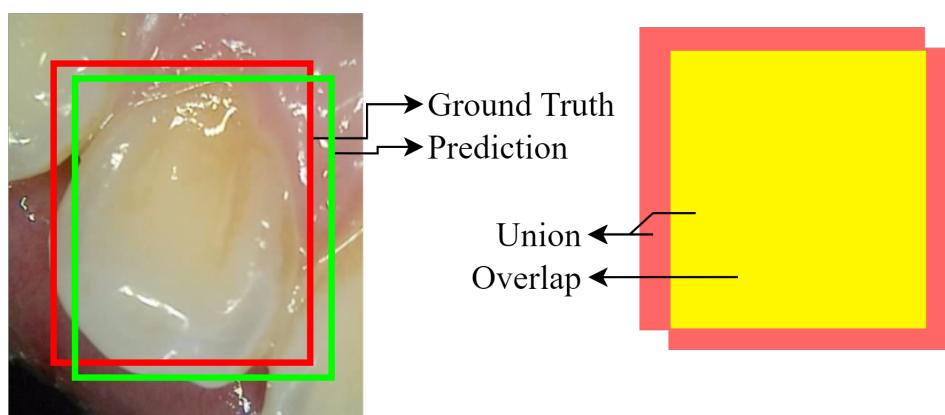
$$F1 - \text{score} = 2 \times \frac{\text{Prec} \times \text{Rec}}{\text{Prec} + \text{Rec}} \quad (3)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (4)$$

*Precision* berfungsi untuk mengukur perbandingan prediksi yang bernilai *true positive* dengan seluruh prediksi positif. *Recall* atau *sensitivity* mendeskripsikan proporsi data yang sesungguhnya bernilai positif dan data yang di prediksi positif. *F1-Score* adalah metrik untuk menilai keseimbangan antara *precision* dan *recall* [17]. mAP adalah metrik khusus yang digunakan pada model deteksi objek untuk mengukur akurasi model. Perhitungan nilai mAP dihasilkan dengan rerata nilai *average precision (AP)* dari setiap kelas [18].

$$IoU = \frac{\text{Area of overlap}}{\text{Area of union}} \quad (5)$$

Gambar 5 menunjukkan bagan dari *Intersection of Union* (IoU) yang digunakan sebagai *threshold* dalam mengukur mAP. IoU adalah nilai yang memutuskan sebuah prediksi bersifat *true positive*, *false negative*, atau *false positive*. Perhitungan nilai IoU didapatkan dari seberapa besar area terprediksi oleh model dengan gabungan area prediksi model dengan area anotasi atau *ground truth* [19].



**Gambar 5.** *Intersection of Union*

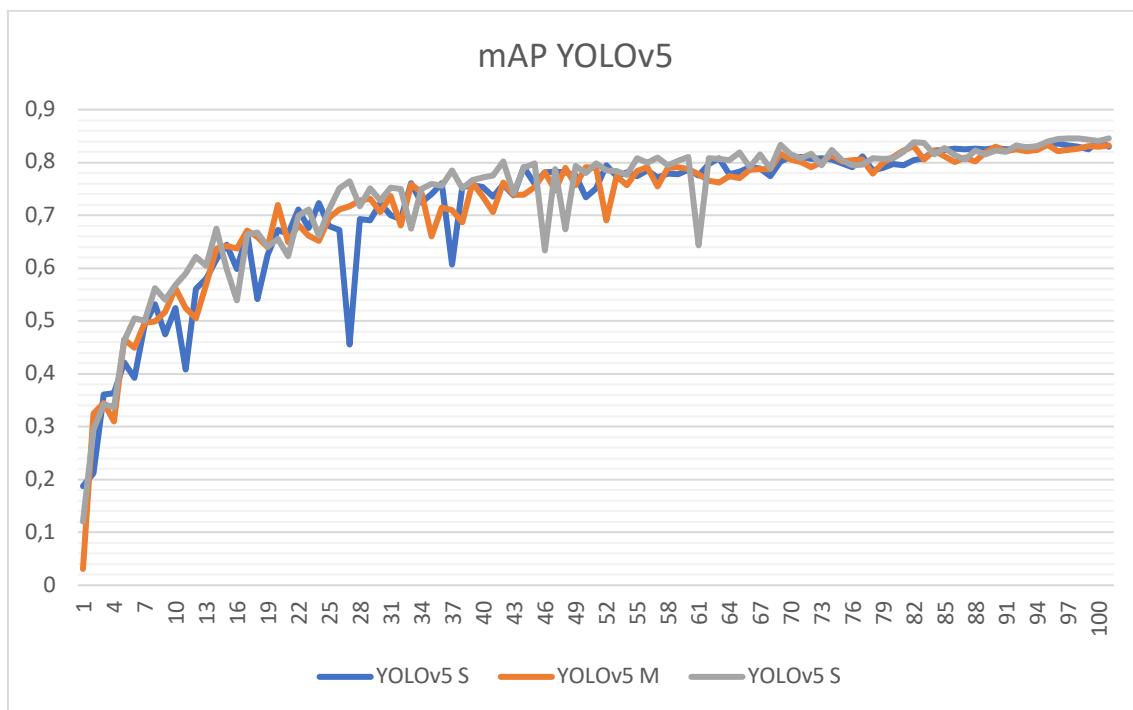
### C. Hasil dan Pembahasan

Proses *training* dan *testing* pada ketiga model diperlakukan dengan parameter dan *resource* yang sama. Pada proses *training*, ukuran data input yang digunakan adalah (224, 224) dengan 100 *epoch*. Google Colaboratory dengan dukungan GPU berjenis T4 digunakan pada proses *training* maupun *testing*. Secara lebih detail parameter yang digunakan pada proses *training* ditunjukkan pada Tabel 2.

**Tabel 2.** Parameter *Training*

Parameter	Value
Input Size	(224, 224)
Epoch	100
Learning Rate (Adam)	0.01
IoU Threshold	0.2

Lama proses *training* berbeda pada masing - masing YOLOv5. Secara arsitektur YOLOv5 S memiliki arsitektur paling sederhana dan YOLOv5 L memiliki arsitektur paling kompleks. Semakin sederhana arsitektur model, semakin cepat dan ringan proses *training*.



**Gambar 6.** Nilai mAP Setiap Epoch

Gambar 6 menunjukkan grafik peningkatan nilai mAP pada setiap *epoch* saat proses *training* dijalankan. Dari grafik tersebut terlihat ketiga model tidak mengalami perkembangan nilai mAP secara signifikan pada *epoch*-85.

Hasil performa ketiga model YOLOv5 dari proses *training* menunjukkan nilai metrik evaluasi yang tidak jauh berbeda. YOLOv5 M dan L mendapatkan nilai mAP tertinggi, namun YOLOv5 L memiliki nilai *precision* yang lebih tinggi. YOLOv5 S hanya memiliki nilai tertinggi pada nilai *recall*. Secara lebih detail hasil *training* pada ketiga model YOLOv5 ditunjukkan pada Tabel 3.

**Tabel 3.** Metrik Evaluasi Hasil *Training*

<b>Model</b>	<b>mAP0.5:0.95</b>	<b>Precision</b>	<b>Recall</b>	<b>F1 Score</b>
YoloV5 S	83%	92%	<b>91%</b>	82%
YoloV5 M	84%	91%	90%	<b>85%</b>
YoloV5 L	<b>84%</b>	<b>93%</b>	87%	84%

Semua output *weight* dari ketiga model kemudian diujicoba menggunakan data *testing*. Proses *testing* dilakukan untuk mengukur performa model dalam menangani data diluar data *training* untuk menghindari terjadinya *overfitting* [20].

**Tabel 4.** Hasil *Testing* Model

<b>Model</b>	<b>Weight Size (MB)</b>	<b>Speed test per image (ms)</b>	<b>Accuracy test</b>
YoloV5 S	13	<b>7 - 12</b>	76%
YoloV5 M	40	10 - 22	<b>82%</b>
YoloV5 L	88	16 - 26	72%

Tabel 4 menunjukkan hasil uji coba pada ketiga model YOLOv5. Dari hasil tersebut YOLOv5 M merupakan model yang terbaik dengan nilai akurasi *testing* sebesar 82%. Di sisi lain, YOLOv5 L yang memiliki arsitektur yang lebih kompleks tidak mampu dengan baik menangani data *testing* dan hanya mendapatkan nilai akurasi *testing* sebesar 72%. Hasil yang cukup baik didapatkan oleh YOLOv5 S dengan nilai akurasi *testing* 76%, mengingat bahwa model tersebut sangat ringan dan memiliki arsitektur yang sederhana.

#### D. Simpulan

Pada penelitian ini metode deteksi objek dimanfaatkan untuk mengidentifikasi jenis gigi pada manusia untuk membantu dokter dalam pemeriksaan gigi menggunakan kamera intraoral. Tiga jenis model YOLOv5 dipilih untuk mengenali lima jenis kondisi gigi yang umum ditemukan pada pasien gigi yaitu : normal, karies, impaksi, sisa akan, dan tumpatan.

Tabel 3 dan 4 menunjukkan bahwa model YOLOv5 dapat dengan baik digunakan untuk mendeteksi dan mengklasifikasi kondisi gigi manusia pada citra intraoral, khusunya model YOLOv5 M yang memiliki performa lebih baik dari ketiga jenis model. Hal ini dapat membuka perkembangan lebih lanjut dalam penelitian ini. Tujuan utama dari penelitian ini adalah untuk menciptakan model *deep learning* yang dapat mengidentifikasi beberapa kondisi gigi. Penelitian ini kemudian dapat digunakan untuk membantu dokter dalam mengidentifikasi gigi dengan cepat dan efisien.

Meskipun model yang digunakan menujukkan performa yang baik, beberapa batasan memengaruhi kinerja model. Batasan terbesar adalah kurangnya dataset yang digunakan. Dataset masih bersifat tidak seimbang antar kelasnya. Pada penggunaan *deep learning*, jumlah data yang digunakan berpengaruh kepada model yang dihasilkan [20]. Batasan ini terjadi karena sulitnya mendapatkan gambar intraoral dari kondisi gigi tertentu. Impaksi dan sisa akar adalah kondisi gigi pada manusia yang cukup jarang ditemui. Untuk mengatasi batasan tersebut dilakukan augmentasi data untuk memperkaya variasi data.

Penelitian selanjutnya dapat berkonsentrasi pada pembuatan dataset dengan rentang data yang lebih luas, untuk menyeimbangkan rasio setiap kelas kondisi gigi.

Data yang lebih banyak diharapkan dapat meningkatkan kinerja model, sehingga dapat dihasilkan model *deep learning* yang lebih robust.

## E. Referensi

- [1] E. D. Fadhillah, P. C. Bramastagiri, R. Sigit, S. Sukaridhoto, A. Brahmanta, and B. S. B. Dewantara, "Smart Odontogram: Dental Diagnosis of Patients Using *Deep learning*," in *2021 International Electronics Symposium (IES)*, 2021, pp. 532–537. doi: 10.1109/IES53407.2021.9594027.
- [2] M. Yusuf, S. Setiawan, and A. Kusuma, "GAMBARAN KELENGKAPAN DAN KESERAGAMAN PENULISAN ODONTOGRAM OLEH DOKTER GIGI DI KOTA SEMARANG," *ODONTO: Dental Journal*, vol. 6, p. 40, Aug. 2019, doi: 10.30659/odj.6.0.40-44.
- [3] J.-H. Lee, D.-H. Kim, S.-N. Jeong, and S.-H. Choi, "Detection and diagnosis of dental caries using a *deep learning*-based *Convolutional Neural Network* algorithm," *J Dent*, vol. 77, pp. 106–111, 2018, doi: <https://doi.org/10.1016/j.jdent.2018.07.015>.
- [4] J.-J. Hwang, Y.-H. Jung, B.-H. Cho, and M.-S. Heo, "An overview of *deep learning* in the field of dentistry," *Imaging Sci Dent*, vol. 49, p. 1, Aug. 2019, doi: 10.5624/isd.2019.49.1.1.
- [5] M. Prados-Privado, J. García Villalón, C. H. Martínez-Martínez, C. Ivorra, and J. C. Prados-Frutos, "Dental Caries Diagnosis and Detection Using Neural Networks: A Systematic Review," *J Clin Med*, vol. 9, no. 11, 2020, doi: 10.3390/jcm9113579.
- [6] L. Lian, T. Zhu, F. Zhu, and H. Zhu, "Deep learning for Caries Detection and Classification," *Diagnostics*, vol. 11, no. 9, 2021, doi: 10.3390/diagnostics11091672.
- [7] K. Moutselos, E. Berdouses, C. Oulis, and I. Maglogiannis, "Recognizing Occlusal Caries in Dental Intraoral Images Using *Deep learning*," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019, pp. 1617–1620. doi: 10.1109/EMBC.2019.8856553.
- [8] J. Kühnisch, O. Meyer, M. Hesenius, R. Hickel, and V. Gruhn, "Caries Detection on Intraoral Images Using *Artificial intelligence*," *J Dent Res*, vol. 101, p. 002203452110325, Aug. 2021, doi: 10.1177/00220345211032524.
- [9] X. Zhang *et al.*, "Development and evaluation of *deep learning* for screening dental caries from oral photographs," *Oral Dis*, vol. 28, no. 1, pp. 173–181, Jan. 2022, doi: 10.1111/odi.13735.
- [10] M. T. G. Thanh, N. Van Toan, V. T. N. Ngoc, N. T. Tra, C. N. Giap, and D. M. Nguyen, "Deep learning Application in Dental Caries Detection Using Intraoral Photos Taken by Smartphones," *Applied Sciences (Switzerland)*, vol. 12, no. 11, Jun. 2022, doi: 10.3390/app12115504.
- [11] Y. E. Almalki *et al.*, "Deep learning Models for Classification of Dental Diseases Using Orthopantomography X-ray OPG Images," *Sensors*, vol. 22, no. 19, 2022, doi: 10.3390/s22197370.
- [12] S. J. Nelson, *Wheeler's dental anatomy, physiology and occlusion-e-book*. Elsevier Health Sciences, 2014.

- [13] M. Mahasin and I. A. Dewi, "Comparison of CSPDarkNet53, CSPResNeXt-50, and EfficientNet-B0 Backbones on YOLO V4 as Object Detector", doi: 10.52088/ijesty.v1i4.291.
- [14] B. Wu, C. Pang, X. Zeng, and X. Hu, "ME-YOLO: Improved YOLOv5 for Detecting Medical Personal Protective Equipment," *Applied Sciences (Switzerland)*, vol. 12, no. 23, Dec. 2022, doi: 10.3390/app122311978.
- [15] M. Horvat, L. Jelečević, and G. Gledec, "A comparative study of YOLOv5 models performance for image localization and classification Hascheck-Croatian Academic Spelling Checker View project A comparative study of YOLOv5 models performance for image localization and classification," 2022. [Online]. Available: <https://www.researchgate.net/publication/363824867>
- [16] D. Snegireva and A. Perkova, "Traffic Sign Recognition Application Using Yolov5 Architecture," in *2021 International Russian Automation Conference (RusAutoCon)*, 2021, pp. 1002–1007. doi: 10.1109/RusAutoCon52004.2021.9537355.
- [17] Ž. Vujović, "Classification Model Evaluation Metrics," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 599–606, 2021, doi: 10.14569/IJACSA.2021.0120670.
- [18] X. Zhu, S. Lyu, X. Wang, and Q. Zhao, "TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios," Aug. 2021, [Online]. Available: <http://arxiv.org/abs/2108.11539>
- [19] S. M. S. Saladin, M. D. S. Ullaa, S. Ahmed, N. Mohammed, T. H. Farook, and J. Dudley, "One-Stage Methods of Computer Vision Object Detection to Classify Carious Lesions from Smartphone Imaging," *Oral*, vol. 3, no. 2, pp. 176–190, Apr. 2023, doi: 10.3390/oral3020016.
- [20] Md. Z. Alom *et al.*, "A State-of-the-Art Survey on Deep learning Theory and Architectures," *Electronics (Basel)*, vol. 8, p. 292, Aug. 2019, doi: 10.3390/electronics8030292.