

Deteksi Ujung Jari menggunakan Faster-RCNN dengan Arsitektur Inception v2 pada Citra Derau

Derry Alamsyah, Teknik Informatika dan Dicky Pratama, Sistem Informasi, STMIK GI MDP

Abstrak—Deteksi ujung jari merupakan bidang pada visi komputer yang memiliki pemanfaatan yang luas dalam hal antarmuka alami, robot, dll. CNN menjadi salah satu metode yang sedang banyak diterapkan dalam deteksi objek, dengan beberapa pembaharuan CNN berevolusi menjadi *Faster-RCNN* yang mampu mendeteksi objek dengan sangat baik. Penelitian ini memanfaatkan kemampuan *Faster-RCNN* dalam mendeteksi ujung jari dengan arsitektur *Inception V2*. Penerapan dilakukan pada citra yang memiliki derau dan tidak memiliki derau. Hasil penelitian menunjukkan citra yang tidak memiliki derau memiliki hasil akurasi deteksi yang baik yaitu 91%, sementara itu untuk masing-masing citra derau: *Gaussian*, *Salt and Pepper*, *Poisson* dan *Speckle* memiliki akurasi sebesar 34%, 5%, 80% dan 21%.

Kata Kunci— Deteksi Ujung Jari, Faster RCNN, Inception V2

Abstract—Fingertip detection is a field on computers that has extensive space in field: NUI, robotics, etc. CNN is one method that is being used in object detection, with some CNN updates being faster - RCNN is able to detect objects very well. This study conducted the ability of *Faster-RCNN* in detecting fingertips with the *Inception V2* architecture. Implementation is done on images that have noise and not. The results showed that image without noise has 91% accuracy, while for each noisy image: *Gaussian*, *Salt and Pepper*, *Poisson* and *Speckle* had an accuracy of 34%, 5%, 80% and 21%.

Index Terms— *fingertip Detection*, *Faster RCNN*, *Inception V2*

I. PENDAHULUAN

Deteksi ujung jari merupakan bagian dari teknologi Antar Muka Alami (*Natural User Interface (NUI)*) pada komputer. Sebagai contoh, pengguna komputer dapat menggunakan tangannya untuk mengoperasikan komputer tanpa menyentuhnya. Dalam hal tersebut diantaranya dapat menggantikan peran *mouse* dengan ujung jari sebagai penunjuk (*pointer*), mengatur tampilan objek lebih besar atau kecil, dsb.

Selain penggunaan pada NUI, deteksi ujung jari juga dapat digunakan untuk mensimulasikan atau menganimasikan gerakan tangan. Hal ini dapat bermanfaat untuk menggerakkan tangan robot guna menyelesaikan pekerjaan yang sulit dilakukannya dengan tangan manusia.

Deteksi ujung jari dapat dicapai melalui operasi morfologi atau pemahaman struktur geometri objek. Beberapa metode terkait hal tersebut diantaranya *convex hull* [1], *gradien* [2]

dan lainnya. Keduanya bertujuan untuk menemukan kurva yang ada pada ujung jari. Pada penelitian [1] pra proses yang dapat digunakan adalah deteksi tepi.

Selain menggunakan metode morfologi atau pemahaman struktur geometri objek, deteksi ujung jari dapat dilakukan dengan mencari fitur-fitur tertentu yang dapat digunakan untuk mengidentifikasi ujung jari. Beberapa metode terkait yang dapat digunakan adalah deteksi lingkaran [3], deteksi lengkungan (*curvature*) [4], atau bahkan hanya menggunakan warna [3]. Pengenalan warna saat ini, dapat digantikan dengan menggunakan citra kedalaman [5][6], dimana informasi kedalaman dapat diperoleh dari alat atau sensor khusus [1]. Salah satu alat yang dapat digunakan adalah *Kinect* [1] atau jika tidak dapat menggunakan metode stokastik dengan memanfaatkan informasi dari setiap pixel secara *time series* [7].

Perkembangan metode deteksi objek saat ini ditunjukkan oleh tren penggunaan metode *deep learning*, yang salah satunya adalah *Convolutional Neural Network (CNN)*. Metode itu sendiri telah berevolusi dari segi arsitektur, yang bertujuan untuk mengurangi kompleksitas pencarian ciri [8]. Selanjutnya penggunaan CNN dapat menjadi dasar untuk metode lain guna mendapatkan keandalan dari metode *deep learning*, yaitu *Faster R-CNN* yang merupakan perbaikan dari *RCNN* dan *Fast RCNN*. Metode *Faster RCNN* menggunakan *Region Proposal Network* [9]. Beberapa contoh penggunaan deteksi objek dengan metode ini adalah mengenali buku [10], mengenali pejalan kaki [11], dll. Dengan memanfaatkan keandalan dari *deep learning*, maka penelitian ini menggunakan *Faster RCNN* dengan Arsitektur *Inception V2* untuk mendeteksi ujung jari pada kasus citra derau.

II. TINJAUAN PUSTAKA

A. Citra Derau

Derau (*Noise*) pada citra merupakan nilai intensitas warna yang tidak diharapkan. Hal tersebut dikarenakan nilai intensitas tersebut dapat membuat efek yang tidak diinginkan seperti artifak, garis yang tidak realistis, garis yang tidak terlihat, sudut, objek yang buram (*blurred object*) dan latar belakang (*background*) yang mengganggu. Sementara itu, Citra Derau merupakan variasi acak dari nilai kecerahan atau informasi pada citra. Derau pada citra dapat diperoleh dari sensor kamera atau pemindai. Beberapa derau yang umum pada citra diantaranya *Gaussian Noise*, *Impulse Value Noise*

(Salt and Pepper Noise), Speckle Noise dan Photon Noise (Poisson Noise) [1].

1. Gaussian noise

Gaussian Noise pada citra dapat dimodelkan, yaitu dengan cara memberikan/menambahkan citra semula dengan derau, yang dikenal dengan Image Independent Noise (Additive Noise Model). Bentuk umum dari Image Independent Noise (Additive Noise Model) memiliki bentuk umum yang ditunjukkan oleh persamaan (1).

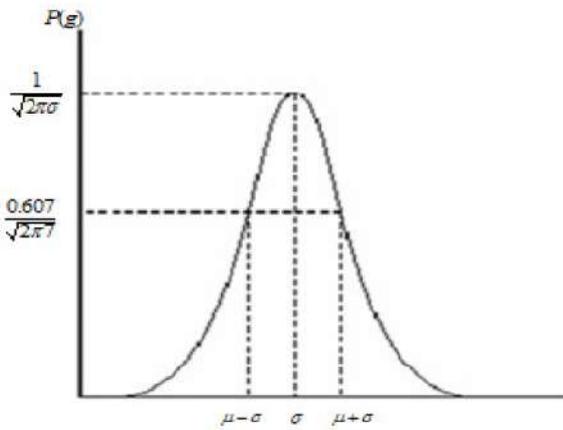
$$I_n = I + f_n \tag{1}$$

Dimana I merupakan citra awal dan f_n merupakan derau (dihasilkan dari model/fungsi distribusi Gauss). Derau yang dihasilkan f_n merupakan nilai acak yang mengikuti distribusi gaussian yang ditunjukkan oleh persamaan (2), sementara itu fungsi distribusi Gauss sendiri ditunjukkan oleh persamaan (3).

$$f_n = \sigma \cdot rand_{gauss} \tag{2}$$

$$P(g) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{g-\mu}{\sigma}\right)^2} \tag{3}$$

Dengan g merupakan nilai intensitas (keabuan) pada citra [12]. Sebagai contoh, citra dengan derajat keabuan 256 dan varians $\sigma^2 = 0.1$, sementara rerata $\mu = 0$ memiliki fungsi distribusi yang ditunjukkan oleh Gambar 1.



Gambar 1. Fungsi Distribusi untuk Gaussian Noise.

2. Impulse Value Noise (Salt and Pepper)

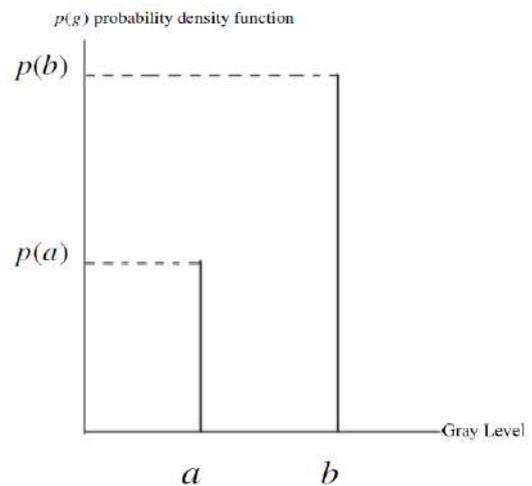
Derau Salt and Pepper dikenal juga sebagai data drop noise karena mengganti nilai dari citra asli. Pixel yang diganti dapat berdekatan (bertetangga) atau tidak, dimana kemungkinannya mengikuti distribusi pada persamaan (4).

$$P(g) = \begin{cases} P_a, & \text{untuk } g = a \\ P_b, & \text{untuk } g = b \\ 0, & \text{selainnya} \end{cases} \tag{4}$$

254	207	210	254	207	210
97	212	32	97	0	32
62	106	20	62	106	20

Gambar 2. Nilai intensitas pusat pixel yang terpengaruh derau

Sebagai contoh, ambil citra 3 x 3 seperti yang ditunjukkan pada Gambar (2) dengan nilai intensitas pusat pixelnya berubah menjadi 0, terpengaruh oleh derau Salt and Pepper. Perubahan yang diberikan oleh derau ini menyisipkan nilai pixel “mati” yaitu 0, dimana nilai tersebut merupakan nilai yang berada pada wilayah gelap, sementara jika nilai tersebut berada pada wilayah terang digantikan dengan nilai paling terang dari citra yang diberikan. Hal tersebut dideskripsikan kedalam fungsi distribusi seperti yang ditunjukkan oleh persamaan (4) dan Gambar 3.



Gambar 3. Fungsi Distribusi untuk Salt and Pepper Noise.

3. Speckle noise

Berbeda dengan Gaussian Noise, Speckle Noise merupakan Jenis derau multiplicative noise. Derau ini dapat terlihat seperti Gaussian Noise, dimana derau yang dihasilkan mengikuti distribusi Gamma (yang ditunjukkan oleh persamaan (5)).

$$P(g) = \frac{g^{\alpha-1} e^{-\frac{g}{a}}}{(\alpha-1)! a^\alpha} \tag{5}$$

4. Photon Noise (Poisson Noise)

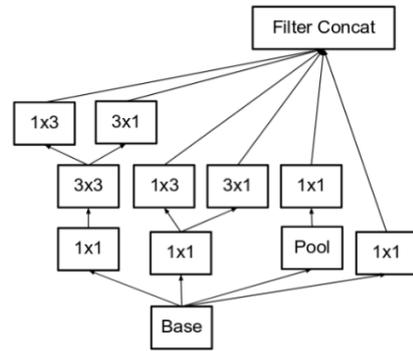
Photon Noise atau yang disebut juga sebagai Poisson Noise merupakan derau yang dihasilkan dari sifat statistik alami dari gelombang elektromagnet seperti gelombang X-ray, Cahaya dan Gamma. Gelombang X-ray dan Gamma memancarkan sejumlah foton (photon) per satuan waktu tertentu. Gelombang tersebut memiliki fluktuasi foton. Selanjutnya, Model dari derau ini mengikuti distribusi Poisson, seperti yang ditunjukkan oleh persamaan (6).

$$P(f_{p_i} = k) = \frac{\lambda_i^k e^{-\lambda}}{k!} \tag{6}$$

B. Faster Region-based Convolutional Neural Network (Faster-RCNN)

Faster RCNN merupakan salah satu metode *deep learning* yang digunakan untuk mengenali suatu objek pada citra. Pengenalan dilakukan dengan menelusuri ciri-ciri yang dimiliki oleh objek pada citra. Penelusuran dilakukan melalui sejumlah *layer* (seperti yang dilakukan pada *neural network*) melalui proses konvolusi atau yang lebih dikenal dengan nama Convolutional Neural Network (CNN).

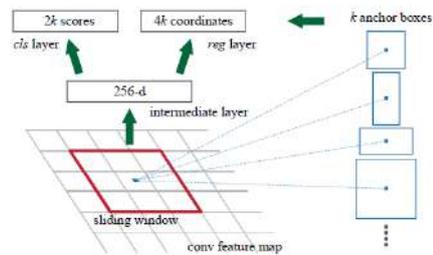
CNN memiliki berbagai arsitektur, salah satunya dalam penelitian ini adalah *Inception V2*. Arsitektur dari *Inception V2* dirancang untuk mengurangi kompleksitas CNN, yang dilakukan dengan cara menyusun arsitektur yang lebih melebar dari pada mendalam. *Inception V2* memiliki 3 modul yang ditunjukkan oleh Gambar 4. Modul pertama (Gambar 4.a) menggantikan konvolusi 5×5 menjadi 3×3 . Selanjutnya pemfaktoran konvolusi dilakukan (ditunjukkan pada Gambar 4.b). Terakhir modul diubah lebih melebar untuk mengurangi kompleksitas jaringan konvolusi (ditunjukkan pada Gambar 4.c) [8].



(c). Modul ketiga *Inception V2*.

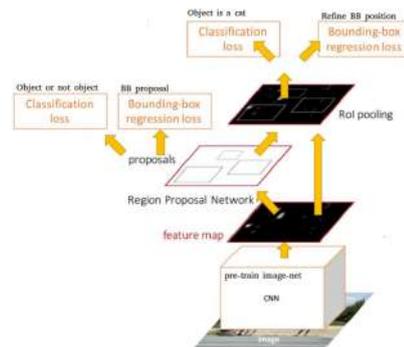
Gambar 4. Modul *Inception V2*.

Selanjutnya, berbeda dengan metode RCNN sebelumnya, maka *Faster-RCNN* membuat perubahan dengan memunculkan *Region Proposal Network* (RPN) (ditunjukkan oleh Gambar 5) [9].

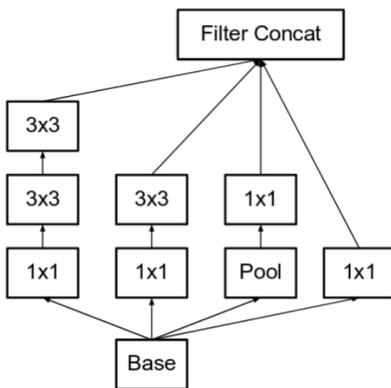


Gambar 5. Region Proposal Network (RPN)

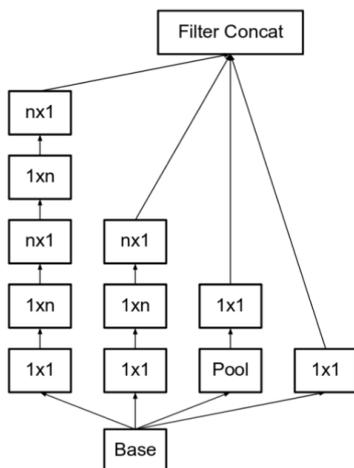
Selanjutnya secara keseluruhan arsitektur dari *Faster RCNN* ditunjukkan oleh Gambar 6.



Gambar 6. Arsitektur Faster R-CNN.



(a). Modul Pertama *Inception V2*.



(b). Modul kedua *Inception V2*.

III. HASIL DAN PEMBAHASAN

Penelitian ini menggunakan data citra telapak tangan yang disediakan oleh IT Departement Mutah University dalam bentuk citra hasil webcam (*Webcam Hand Image Database* (WEHI)). Citra yang digunakan adalah 40 citra, dimana komposisi dataset dibagi menjadi masing-masing 20 citra untuk tahap pelatihan (*Training Phase*) dan *Testing Phase*). Dataset yang digunakan ditunjukkan pada Gambar 7.



Gambar 7. Dataset Telapak Tangan.

Tahap Pelatihan F-RCNN dilakukan terhadap data citra tanpa derau, sementara itu untuk tahap pengujian dilakukan pada citra dengan dan tanpa derau. Hasil pengujian ditunjukkan oleh Tabel 1 dan 2.

Tabel 1. Hasil Deteksi

Citra	Tanpa Derau	Gaussian	Salt & Peper	Poisson	Speckle
1	5	2	0	5	0
2	5	3	0	5	2
3	5	3	0	5	4
4	1	0	0	0	0
5	4	1	0	3	0
6	4	0	0	1	0
7	5	1	0	3	0
8	5	0	0	5	0
9	5	0	0	5	0
10	5	0	0	5	0
11	5	3	1	5	1
12	5	2	0	5	2
13	4	2	0	4	2
14	5	5	2	5	4
15	4	2	0	4	1
16	5	3	0	5	1
17	4	0	0	4	0
18	5	0	0	2	1
19	5	2	0	4	1
20	5	5	2	5	2

Secara spesifik, masing-masing parameter untuk derau adalah:

1. Parameter untuk *Gaussian Noise*: rata rata = 0 dan Varians 0.01.
2. Parameter untuk *Salt and Pepper Noise*: noise density = 0.05 (mempengaruhi 5% pixel).
3. Parameter untuk *Speckle Noise*: rata-rata = 0 dan varians = 0.04.

Tabel 2. Hasil Deteksi Ujung Jari

Hasil	Keterangan
	Tanpa Derau
	Gaussian
	Salt & Pepper
	Poisson
	Speckle

Berdasarkan Tabel 3 deteksi mendapatkan hasil akurasi yang paling baik adalah pada citra tanpa derau, sementara untuk citra derau hasil yang paling baik didapat pada citra derau Poisson. Sementara itu hasil yang kurang baik didapat pada citra derau *Speckle* dan *Gaussian*, serta yang paling buruk didapat dari citra *salt and pepper*.

Tabel 3. Hasil Akurasi

Tanpa Derau	Gaussian	Salt & Peper	Poisson	Speckle
91%	34%	5%	80%	21%

IV. KESIMPULAN

Kemampuan Faster-RCNN untuk mendeteksi ujung jari dalam kondisi normal (tanpa derau) sangat baik, yaitu memiliki akurasi 91%. Akan tetapi, untuk derau Salt & Pepper, hasil akurasi deteksi hanya 5%. Derau yang mampu diatasi dengan baik oleh Faster-RCNN adalah derau Poisson, dimana akurasi mencapai 80%. Untuk derau Gaussian dan Speckle, Faster-RCNN masih belum cukup baik untuk mengatasi dengan nilai akurasi 34% dan 21%.

REFERENSI

- [1] D. Alamsyah and M. I. Fanany. 2013. Particle filter for 3D fingertips tracking from color and depth images with occlusion handling. IEEE conference on International Conference on Advanced Computer Science and Information Systems (ICACSIS). pp. 445-449.
- [2] A. Wang, H. Lu, and H. Lu. 2015. An effective real-time fingertip positioning system based on gradient information extraction from frame image sequences. 8th International Congress on Image and Signal Processing (CISP). pp. 330-334.
- [3] Md. J. Alam and M. Chowdhury. 2013. Detection of fingertips based on the combination of color information and circle detection. IEEE 8th International Conference on Industrial and Information Systems. pp. 572 - 576.
- [4] G. Wu and W. Kang. 2017. Vision-Based Fingertip Tracking Utilizing Curvature Points Clustering and Hash Model Representation. IEEE Transactions on Multimedia. pp. 1730-1741. Vol. 19.
- [5] Y. Bak, M. Li, L. Sun, and Q. Huo. 2017. Fingertip detection based on protuberant saliency from depth image. IEEE International Conference on Image Processing (ICIP). pp. 3380-3384.
- [6] C. Liang, Y. Song, and Y. Zhang. 2015. Real-time fingertip detection based on depth data. 3rd IAPR Asian Conference on Pattern Recognition (ACPR).
- [7] M. Rachmadi and D. Alamsyah. 2017. Estimasi Citra Kedalaman Dengan Conditional Random Field (CRF) dan Structured Support Vector Machine (SSVM). Jurnal Rekayasa Sistem dan Teknologi Informasi (RESTI). pp. 198-203. Vol. 1. No. 3.
- [8] C. Szegedy, V. Vanhoucke, S. Ioffe, and J. Shlens. 2016. Rethinking the Inception Architecture for Computer Vision. IEEE conference on Computer Vision and Pattern Recognition (CVPR). pp. 2818-2826.
- [9] S. Ren, K. He, R. Girshick, and J. Sun. 2016. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE transactions on Pattern Analysis and Machine Intelligence. pp. 1137-1149. Vol. 39.
- [10] B. Zhu, X. Wu, L. Yang, Y. Shen, and L. Wu. 2016. *Automatic detection of books based on Faster R-CNN.*

Third International Conference on Digital Information Processing, Data Mining, and Wireless Communication (DIPDMWC). pp. 8-12.

- [11] H. Zhang, Y. Du, S. Ning, Y. Zhang, S. Yang and C. Du. 2017. *Pedestrian Detection Method Based on Faster R-CNN.* 13th International Conference on Computational Intelligence and Security (CIS). pp. 427-430.
- [12] A. K. Boyat and B. K. Joshi. *A Review Paper: Noise in Digital Image Processing.* Signal & Image Processing: An International Journal (SIPIJ). Vol. 6, No. 2. 2015. pp 63-75.



Derry Alamsyah, S.Si, M.Kom, M.Pd lahir di Jakarta pada tanggal 22 Januari 1989. Penulis mendapatkan gelar Sarjana Sains (S.Si) pada tahun 2009 dari FMIPA Matematika Universitas Sriwijaya dan Komputer (S.Kom) pada tahun 2012 dari program studi Teknik Informatika, STMIK GI MDP. Kemudian penulis melanjutkan

pendidikan S2 pada Fakultas Ilmu Komputer, Magister Ilmu Komputer, Universitas Indonesia, dan telah menyelesaikan jenjang pendidikan master pada tahun 2013 begitupun pada FKIP, Magister Pendidikan Matematika, Universitas Sriwijaya pada tahun 2014. Adapun bidang ilmu yang ditekuni penulis adalah pengolahan citra dan pembelajaran mesin.



Dicky Pratama, S.Kom., M.T.I lahir di Palembang pada tanggal 11 Nopember 1986. Penulis mendapatkan gelar Sarjana Komputer (S.Kom) pada tahun 2011 dari program studi Teknik Informatika, STMIK GI MDP. Kemudian penulis melanjutkan pendidikan S2 pada Fakultas Ilmu Komputer, Jurusan Magister Teknologi

Informasi, Universitas Indonesia, dan telah menyelesaikan jenjang pendidikan master pada tahun 2014. Adapun bidang ilmu yang ditekuni penulis adalah Sistem Informasi dan Audit Sistem Informasi.