

***EXEMPLAR BASED CONVOLUTIONAL NEURAL NETWORK UNTUK
PENCARIAN WAJAH PADA VIDEO REKAMAN CCTV***

SKRIPSI

Diajukan untuk Memenuhi Sebagian dari
Syarat Memperoleh Gelar Sarjana Komputer
Program Studi Ilmu Komputer



Oleh
Winda Mauli Kristy
NIM 1606216

PROGRAM STUDI ILMU KOMPUTER
DEPARTEMEN PENDIDIKAN ILMU KOMPUTER
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PENDIDIKAN INDONESIA
2020

***EXEMPLAR BASED CONVOLUTIONAL NEURAL NETWORK UNTUK
PENCARIAN WAJAH PADA VIDEO REKAMAN CCTV***

Oleh
Winda Mauli Kristy
NIM 1606216

Sebuah Skripsi yang Diajukan untuk Memenuhi Salah Satu Syarat Memperoleh
Gelar Sarjana Komputer di Fakultas Pendidikan Matematika dan Ilmu
Pengetahuan Alam

© Winda Mauli Kristy 2020
Universitas Pendidikan Indonesia
Agustus 2020

Hak Cipta Dilindungi Undang-Undang
Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian, dengan dicetak
ulang, difoto kopi, atau cara lainnya tanpa izin dari penulis

Winda Mauli Kristy

1606216

***EXEMPLAR BASED CONVOLUTIONAL NEURAL NETWORK UNTUK
PENCARIAN WAJAH PADA VIDEO REKAMAN CCTV***

DISETUJUI DAN DISAHKAN OLEH PEMBIMBING:

Pembimbing I,



Yaya Wihardi, M.Kom.

NIP. 198903252015041001

Pembimbing II,



Erlangga, S.Kom., M.T.

NIP. 198607082018031001

Mengetahui,

Kepala Departemen Pendidikan Ilmu Komputer



Lala Septem Riza, M.T., Ph.D.

NIP. 197809262008121001

**EXEMPLAR BASED CONVOLUTIONAL NEURAL NETWORK UNTUK
PENCARIAN WAJAH PADA VIDEO REKAMAN CCTV**

Oleh

Winda Mauli Kristy – windamkristy@gmail.com

1606216

ABSTRAK

Banyak metode yang dapat melakukan pencarian wajah dengan baik, namun pada umumnya metode tersebut membutuhkan banyak sampel, khususnya pencarian wajah dengan menggunakan metode-metode *deep learning*. Akan tetapi, terkadang untuk melakukan pencarian wajah terdapat pula kasus di mana sampel yang didapat hanya berasal dari video rekaman CCTV tersebut, dan tidak memungkinkan melakukan proses *training* terlebih dahulu. Sehingga pendekatan yang diambil haruslah dengan menggunakan pendekatan *exemplar-based*. Pada penelitian ini menggunakan salah satu metode *convolution neural network* (CNN) yang digabungkan dengan dua metode *matching* berbeda, yaitu *cross-correlation matching* (CCM) dan *normalized cross-correlation matching* (NCC). Penelitian dilakukan dengan menggunakan Chokepoint Face Dataset. Di mana data akan dilatih dengan menggunakan *triplet loss optimization*. Penelitian ini pun bertujuan untuk mengetahui kinerja dari penggabungan metode-metode tersebut. Pada setiap metodenya akan dibuat dua arsitektur berbeda dan diuji untuk dilihat hasil akurasi dari masing-masing arsitektur dari kedua metode tersebut. Dari hasil pengujian, metode CNN-NCC memiliki hasil akurasi, yaitu sekitar 2 hingga 17.9% lebih tinggi dari metode CNN-CCM. Namun akurasi yang didapatkan sangat bergantung pada variasi yang ada pada video rekaman CCTV tersebut.

Kata Kunci: Pencarian Wajah, CCTV, *Exemplar-Based, Convolution Neural Network, Cross-correlation Matching, Normalized Cross-correlation Matching, Triplet Loss Optimization*.

**EXEMPLAR BASED CONVOLUTIONAL NEURAL NETWORK FOR FACE
SEARCH ON CCTV RECORD VIDEOS**

Arranged by

Winda Mauli Kristy – windamkristy@gmail.com

1606216

ABSTRACT

Many methods can do a good face search, but in general these methods require a lot of samples, especially face searches using deep learning methods. Will still, sometimes to do a face search there are also cases where the sample obtained only comes from the CCTV video footage, and it is not possible to do the training process first. So the approach taken must use an exemplar-based approach. In this study, one of the methods of convolution neural network (CNN) combined with two different matching methods, namely cross-correlation matching (CCM) and normalized cross-correlation matching (NCC). The study was conducted using Chokepoint Face Dataset. Where the data will be trained using triplet loss optimization. This research also aims to determine the performance of the combination of these methods. For each method, two different architectures will be made and tested to see the accuracy of each architecture from the two methods. From the test results, the CNN-NCC method has an accuracy result, which is about 2 to 17.9% higher than the CNN-CCM method. However, the accuracy obtained really depends on the variations in the CCTV record videos.

Keywords: Face Search, CCTV, Exemplar-Based, Convolution Neural Network, Cross-correlation Matching, Normalized Cross-correlation Matching, Triplet Loss Optimization.

DAFTAR ISI

KATA PENGANTAR	i
UCAPAN TERIMAKASIH.....	ii
ABSTRAK	iv
<i>ABSTRACT</i>	v
DAFTAR ISI.....	vi
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	4
1.4 Batasan Masalah.....	5
1.5 Manfaat Penelitian.....	5
1.6 Sistematika	5
BAB II KAJIAN PUSTAKA	7
2.1 Peta Literatur	7
2.2 Penelitian Terkait.....	8
2.3 Pencarian Wajah.....	10
2.4 Video Rekaman CCTV.....	11
2.5 Python.....	12
2.6 Pengolahan Citra Digital	13
2.6.1 Operasi Pengolahan Citra	14
2.7 <i>Computer Vision</i>	15
2.8 Jaringan Saraf Tiruan	16
2.9 <i>Deep Learning</i>	17

2.10	<i>Face Recognition</i>	17
2.11	<i>Convolutional Neural Network</i>	18
2.10.1	Struktur <i>Convolutional Neural Network</i>	19
2.10.2	Fungsi Aktivasi	21
2.12	<i>Exemplar Based</i>	22
2.13	<i>Triplet Loss Optimization</i>	24
2.14	<i>Cross-correlation Matching</i>	24
2.15	<i>Normalized Cross-correlation Matching</i>	26
	BAB III METODE PENELITIAN.....	28
3.1	Tahapan Penelitian	28
3.1.1	Perumusan Masalah.....	28
3.1.2	Studi Literatur.....	29
3.1.3	Pengumpulan Data.....	29
3.1.4	Rancangan Metode	32
3.1.5	Pengujian	35
3.1.6	Analisis dan Pembahasan	41
3.1.7	Penarikan Kesimpulan.....	41
3.2	Lingkungan Komputasi	41
	BAB IV HASIL DAN PEMBAHASAN	42
4.1	Pengolahan Data.....	42
4.2	Arsitektur Model Pencarian Wajah	43
4.2.1	<i>Convolutional Neural Network</i>	44
4.2.2	<i>Cross-correlation Matching</i>	48
4.2.3	<i>Normalized Cross-correlation Matching</i>	51
4.3	<i>Training</i>	53
4.4	Pengujian	56

4.5	Hasil Pengujian.....	58
4.6	Analisis dan Pembahasan Hasil.....	64
BAB V KESIMPULAN DAN SARAN.....		70
5.1	Kesimpulan.....	70
5.2	Saran	71
DAFTAR PUSTAKA		72

DAFTAR PUSTAKA

- Ahmad, A. (2017). Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning. *Jurnal Teknologi Indonesia*, (October), 3.
- Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2018). Understanding of a convolutional neural network. *Proceedings of 2017 International Conference on Engineering and Technology, ICET 2017, 2018-Janua*(August), 1–6. <https://doi.org/10.1109/ICEngTechnol.2017.8308186>
- Andrej Krenker, Bešter, J., & Kos, A. (2011). Introduction to the Artificial Neural Networks, In: Suzuki K (ed), Artificial Neural Networks: Methodological Advances and Biomedical Applications. *InTech*, 1–18. <https://doi.org/10.5772/15751>
- Arlando Saragih, R. (2007). Pengenalan Wajah Menggunakan Metode Fisherface. *Jurnal Teknik Elektro*, 7(1), 50–61. <https://doi.org/10.9744/jte.7.1.50-62>
- Arrofiqoh, E. N., & Harintaka, H. (2018). Implementasi Metode Convolutional Neural Network Untuk Klasifikasi Tanaman Pada Citra Resolusi Tinggi. *Geomatika*, 24(2), 61. <https://doi.org/10.24895/jig.2018.24-2.810>
- Arsenov, A., Ruban, I., Smelyakov, K., & Chupryna, A. (2018). Evolution of convolutional neural network architecture in image classification problems. *CEUR Workshop Proceedings*, 2318, 35–45.
- Ashby, M. P. J. (2017). The Value of CCTV Surveillance Cameras as an Investigative Tool: An Empirical Analysis. *European Journal on Criminal Policy and Research*. <https://doi.org/10.1007/s10610-017-9341-6>
- Asror, I., & Siraj, Y. (2016). Desain dan Implementasi Cloud Design and Implementation CCTV on Cloud. *Telekontran*, 4(May).
- Bautista, M. A., Sanakoyeu, A., Sutter, E., & Ommer, B. (2016). CliqueCNN: Deep

- unsupervised exemplar learning. *Advances in Neural Information Processing Systems*, (Nips), 3853–3861.
- Cai, B., Jiang, Z., Zhang, H., Yao, Y., & Nie, S. (2018). Online Exemplar-Based Fully Convolutional Network for Aircraft Detection in Remote Sensing Images. *IEEE Geoscience and Remote Sensing Letters*, 15(7), 1095–1099. <https://doi.org/10.1109/LGRS.2018.2829147>
- Chen, S., Liu, J., & Zhou, Z.-H. (2004). Making FLDA applicable to face recognition with one sample per person. *Pattern Recognition*, 37, 1553–1555. [https://doi.org/10.1016/S0031-3203\(04\)00012-3](https://doi.org/10.1016/S0031-3203(04)00012-3)
- Das, K., & Behera, R. N. (2017). A Survey on Machine Learning: Concept, Algorithms and Applications. *International Journal of Innovative Research in Computer and Communication Engineering*, 5(2). <https://doi.org/10.15680/IJIRCCE.2017>
- Demir, H. S., & Akagündüz, E. (2020). Filter design for small target detection on infrared imagery using normalized-cross-correlation layer. *Turkish Journal of Electrical Engineering and Computer Sciences*, 28(1), 302–317. <https://doi.org/10.3906/elk-1807-287>
- Dunn, J. (2018). *Robust search templates for face-in-crowd search James Daniel Dunn A thesis in fulfilment of the requirements for the degree of Doctor of Philosophy School of Psychology Faculty of Science November 2018 Thesis / Dissertation Sheet*. (November).
- Eka Putra, W. S. (2016). Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101. *Jurnal Teknik ITS*, 5(1). <https://doi.org/10.12962/j23373539.v5i1.15696>
- Feng, Z., Qingming, H., & Wen, G. (2006). Image matching by normalized cross-correlation. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2(June 2006).

- <https://doi.org/10.1109/icassp.2006.1660446>
- Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36(4), 193–202. <https://doi.org/10.1007/BF00344251>
- Gao, C., Chen, F., Yu, J. G., Huang, R., & Sang, N. (2017). Robust Visual Tracking Using Exemplar-Based Detectors. *IEEE Transactions on Circuits and Systems for Video Technology*, 27(2), 300–312. <https://doi.org/10.1109/TCSVT.2015.2513700>
- Günay, M. (2017). *Comparison of Face Recognition Algorithms*. 0–3.
- Hu, C., Ye, M., Ji, S., Zeng, W., & Lu, X. (2015). A new face recognition method based on image decomposition for single sample per person problem. *Neurocomputing*, 160, 287–299. <https://doi.org/10.1016/j.neucom.2015.02.032>
- Jian Huang, Yuen, P. C., Wen-Sheng Chen, & Lai, J. H. (2004). Component-based LDA method for face recognition with one training sample. 120–126. <https://doi.org/10.1109/amfg.2003.1240833>
- Kadir, A., & Susanto, A. (2013). *Teori dan Aplikasi Pengolahan Citra*. (January 2013), 640.
- Kanata, B. (2015). Pencocokan Citra Sidik Jari Menggunakan Korelasi Silang Ternormalisasi. *Jurnal Rekayasa Elektrika*, 11(4), 144. <https://doi.org/10.17529/jre.v11i4.2405>
- Katole, A., Yellapragad, K., Bedi, A., Kalra, S., & Chaitanya, M. (2015). Hierarchical Deep Learning Architecture for 10K Objects Classification. 1–134. <https://doi.org/10.1227/01.NEU.0000255452.20602.C9>
- Kerr, C. (2009). *CLOSED CIRCUIT TELEVISION (CCTV) AND ITS APPLICATION TO PUBLIC SAFETY INITIATIVES: Effectiveness and*

Considerations By.

- Koch, G., Zemel, R., & Salakhutdinov, R. (2015). Siamese Neural Networks for One-shot Image Recognition Gregory. *Proceedings of the 32nd International Conference on Machine Learning*, 37. <https://doi.org/10.1136/bmj.2.5108.1355-c>
- Krishna, R. (2017). *Computer Vision: Foundation and Application*. 213.
- Kusumanto, R., & Tompunu, A. N. (2011). Pengolahan Citra Digital Untuk Mendeteksi Objek Menggunakan Pengolahan Warna Model Normalisasi RGB. *Studies in Environmental Science*. [https://doi.org/10.1016/S0166-1116\(08\)71924-1](https://doi.org/10.1016/S0166-1116(08)71924-1)
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Loderer, M., Pavlovicova, J., Oravec, M., & Mazanec, J. (2015). Face parts importance in face and expression recognition. *2015 22nd International Conference on Systems, Signals and Image Processing - Proceedings of IWSSIP 2015*, 188–191. <https://doi.org/10.1109/IWSSIP.2015.7314208>
- Malmenbratt, J., Brooks, D. J., Malmenbratt, J., & Brooks, D. J. (2015). *CCTV surveillance: The differing aims and functions of CCTV within the corporate stratum*. 2015, 28–35. <https://doi.org/10.4225/75/57a83a1bd2cf6>
- Mileva, M., & Burton, A. M. (2019). Face search in CCTV surveillance. *Cognitive Research: Principles and Implications*, 4(1). <https://doi.org/10.1186/s41235-019-0193-0>
- Munir, R. (2013). Pengantar Pengolahan Citra. *Pengolahan Citra Digital*, (Bagian 1), 1–10. Diambil dari <http://rosni-gj.staff.gunadarma.ac.id/Downloads/files/15431/pendahuluan.pdf>
- Nourain Dawoud, N., Belhaouari Samir, B., & Janier, J. (2011). Fast Template

- Matching Method based Optimized Sum of Absolute Difference Algorithm for Face Localization. *International Journal of Computer Applications*, 18(8), 30–34. <https://doi.org/10.5120/2302-2912>
- Nwankpa, C., Ijomah, W., Gachagan, A., & Marshall, S. (2018). *Activation Functions: Comparison of trends in Practice and Research for Deep Learning*. 1–20. Diambil dari <http://arxiv.org/abs/1811.03378>
- Parchami, M., Bashbaghi, S., & Granger, E. (2017). CNNs with cross-correlation matching for face recognition in video surveillance using a single training sample per person. *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2017*, (August). <https://doi.org/10.1109/AVSS.2017.8078554>
- Pegi, I. M., Amerta, K., & Wibawa, I. G. A. (2020). *Alphabet Writing Game Application using Template Matching Cross-correlation*. 8(4), 393–399.
- Prabowo, D., Abdullah, D., & Manik, A. (2018). Deteksi Perhitungan Objek Berdasarkan Warna Menggunakan COLOR Object Trackinng. *Generation Journal*, 2(1), 40–47.
- Salamun, S., & Wazir, F. (2016). Rancang Bangun Sistem Pengenalan Wajah Dengan Metode Principal Component Analysis. *Rabit : Jurnal Teknologi dan Sistem Informasi Univrab*, 1(2), 59–75. <https://doi.org/10.36341/rabit.v1i2.25>
- Sayem, I. M., & Chowdhury, M. S. (2019). Integrating face recognition security system with the internet of things. *Proceedings - International Conference on Machine Learning and Data Engineering, iCMLDE 2018*, 19–21. <https://doi.org/10.1109/iCMLDE.2018.00013>
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 07-12-June*, 815–823. <https://doi.org/10.1109/CVPR.2015.7298682>

- Subramaniam, A., & Chatterjee, M. (2016). Deep Neural Networks with Inexact Matching for Person Re-Identification. *30th Conference on Neural Information Processing Systems*, (1). <https://doi.org/10.1111/j.1540-6261.1974.tb00036.x>
- Suka Parwita, W. G., & Winarko, E. (2015). Hybrid Recommendation System Memanfaatkan Penggalian Frequent Itemset dan Perbandingan Keyword. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 9(2), 167. <https://doi.org/10.22146/ijccs.7545>
- Susanto, B., Purnomo, F., & Fahmi, M. (2017). Sistem Keamanan Pintu Berbasis Pengenalan Wajah Menggunakan Metode Fisherface. *Jurnal Ilmiah Inovasi*, 17. <https://doi.org/10.25047/jii.v17i1.464>
- Szeliski, R. (2010). Computer Vision: Algorithms and Applications. *European Journal of Immunology*. <https://doi.org/10.1002/eji.200526121>
- The U.S. Department of Homeland Security. (2013). *CCTV Technology Handbook*. (July), 66.
- Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of Face Perception. *Annual Review of Neuroscience*, 31(1), 411–437. <https://doi.org/10.1146/annurev.neuro.30.051606.094238>
- Wang, D., Otto, C., & Jain, A. K. (2017). Face Search at Scale. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1122–1136. <https://doi.org/10.1109/TPAMI.2016.2582166>
- Wiley, V., & Lucas, T. (2018). Computer Vision and Image Processing: A Paper Review. *International Journal of Artificial Intelligence Research*, 2(1), 22. <https://doi.org/10.29099/ijair.v2i1.42>
- Wu, J., & Zhou, Z.-H. (2002). Face recognition with one training image per person. *Pattern Recognition Letters*, 23, 1711–1719. [https://doi.org/10.1016/S0167-8611\(01\)00111-7](https://doi.org/10.1016/S0167-8611(01)00111-7)

8655(02)00134-4

Yogitha, S., & Sakthivel, P. (2014). A distributed computer machine vision system for automated inspection and grading of fruits. *Proceeding of the IEEE International Conference on Green Computing, Communication and Electrical Engineering, ICGCCEE 2014*, 3(4), 522–527.
<https://doi.org/10.1109/ICGCCEE.2014.6922281>

Zhi, T., Duan, L. Y., Wang, Y., & Huang, T. (2016). Two-stage pooling of deep convolutional features for image retrieval. *Proceedings - International Conference on Image Processing, ICIP, 2016-Augus*, 2465–2469.
<https://doi.org/10.1109/ICIP.2016.7532802>

Zou, F., Yang, F., Chen, W., Li, K., Song, J., Chen, J., & Ling, H. (2019). Fast large scale deep face search. *Pattern Recognition Letters*.
<https://doi.org/10.1016/j.patrec.2019.01.012>

Zufar, M., & Setiyono, B. (2016). Convolutional Neural Networks Untuk Pengenalan Wajah Secara Real-Time. *Jurnal Sains dan Seni ITS*, 5(2), 72–77.
<https://doi.org/10.12962/j23373520.v5i2.18854>