

UNIVERSITAS GUNADARMA
FAKULTAS TEKNOLOGI INDUSTRI



**ESTIMASI POSE TIGA DIMENSI DARI GAMBAR
MONOKULER MENGGUNAKAN DEEP NEURAL
NETWORK**

Disusun oleh:

Nama	: Denilson
NPM	: 51416815
Jurusan	: Teknik Informatika
Pembimbing	: Dr. Dharmayanti, ST., MMSI.

**Diajukan Guna Melengkapi Sebagian Syarat
Dalam Mencapai Gelar Sarjana Strata Satu (S1)**

Depok

2020

LEMBAR ORIGINALITAS DAN PUBLIKASI

Saya yang bertanda tangan di bawah ini,

Nama : Denilson
NPM : 51416815
Judul Penulisan Ilmiah : Estimasi Pose Tiga Dimensi dari Gambar Monokuler
Menggunakan Deep Neural Network
Tanggal Sidang : tanggal
Tanggal Lulus : tanggal

menyatakan bahwa tulisan ini adalah merupakan hasil karya saya sendiri dan dapat dipublikasikan sepenuhnya oleh Universitas Gunadarma. Segala kutipan dalam bentuk apa pun telah mengikuti kaidah, etika yang berlaku. Mengenai isi dan tulisan adalah merupakan tanggung jawab Penulis, bukan Universitas Gunadarma.

Demikian pernyataan ini dibuat dengan sebenarnya dan dengan penuh kesadaran.

Depok, April 2020

Denilson

LEMBAR PENGESAHAN

Komisi Pembimbing

No	Nama	Kedudukan
1	Dr. Dharmayanti, ST., MMSI.	Ketua
2	DIGANTI NAMA PENGUJI 2	DIGANTI JABATAN PENGUJI 2
3	DIGANTI NAMA PENGUJI 3	DIGANTI JABATAN PENGUJI 3

Tanggal Sidang : tgl bln thn

Panitia Ujian

No	Nama	Kedudukan
1	DIGANTI NAMA PENGUJI 1	DIGANTI JABATAN PENGUJI 1
2	DIGANTI NAMA PENGUJI 2	DIGANTI JABATAN PENGUJI 2
3	DIGANTI NAMA PENGUJI 3	DIGANTI JABATAN PENGUJI 3
4	DIGANTI NAMA PENGUJI 4	DIGANTI JABATAN PENGUJI 4
5	DIGANTI NAMA PENGUJI 5	DIGANTI JABATAN PENGUJI 5

Tanggal Lulus : tgl bln thn

MENGETAHUI

Pembimbing

Bagian Sidang Sarjana

(Dr. Dharmayanti, ST., MMSI.)

(NAMA BAGIAN SARJANA)

ABSTRAKSI

Denilson, 51416815

ESTIMASI POSE TIGA DIMENSI DARI GAMBAR MONOKULER
MENGUNAKAN DEEP NEURAL NETWORK

Tugas Akhir. Jurusan Teknik Informatika, Fakultas Teknologi Industri,
Universitas Gunadarma, 2020

Kata Kunci : dibuaturut abjad sekitar 3-5 kata kunci
(jml hlm romawi + jml hlm arab + Lampiran)

Abstraksi.

Daftar Pustaka (thn terlama-thn terbaru)

ABSTRACT

Denilson, 51416815

THREE DIMENSIONAL POSE ESTIMATION FROM MONOCULAR IMAGE
USING DEEP NEURAL NETWORK

Thesis. Informatics Engineering, Faculty of Industrial Technology,
Gunadarma University, 2020

Keywords: dibuaturut abjad sekitar 3-5 kata kunci
(jml hlm romawi + jml hlm arab + Lampiran)

Abstract.

Bibliography (thn terlama-thn terbaru)

KATA PENGANTAR

Segala puji dan syukur penulis ucapkan ke hadirat Tuhan Yang Maha Esa yang telah memberikan berkat, anugerah dan karunia yang melimpah, sehingga penulis dapat menyelesaikan Tugas Akhir ini pada waktu yang telah ditentukan.

Tugas Akhir ini disusun guna melengkapi sebagian syarat untuk memperoleh gelar Sarjana Teknik Informatika Universitas Gunadarma. Adapun judul Tugas Akhir ini adalah "Estimasi Pose Tiga Dimensi Dari Gambar Monokuler Menggunakan Deep Neural Network".

Walaupun banyak kesulitan yang penulis harus hadapi ketika menyusun Tugas Akhir ini, namun berkat bantuan dan dorongan dari berbagai pihak, akhirnya Tugas Akhir ini dapat diselesaikan dengan baik. Untuk itu penulis tidak lupa mengucapkan terima kasih kepada:

1. Ibu Prof. E. S. Margianti, SE, MM selaku rektor Universitas Gunadarma
2. selaku Dekan Fakultas Universitas Gunadarma
3. selaku Ketua Jurusan
4. selaku Bagian Sidang Sarjana
5. Ibu Dr. Dharmayanti, ST., MMSI sebagai pembimbing penulis yang ditengah-tengah kesibukannya telah membimbing penulis sehingga penulisan ini dapat diselesaikan.
6. Keluarga yang selalu mendukung dan terus memberikan motivasi.
7. Semua pihak yang terlibat dalam membantu penyelesaian Tugas Akhir ini.

Sebagai manusia biasa yang tak luput dari kesalahan, maka penulis meminta maaf atas segala kekurangan dan keterbatasan dalam penyusunan Tugas Akhir ini. Penulis sadari bahwa penulisan ini masih jauh dari sempurna, disebabkan karena berbagai keterbatasan yang penulis miliki. Untuk itu penulis mengharapkan kritik

dan saran yang bersifat membangun untuk menjadi perbaikan di masa yang akan datang.

Akhir kata, penulis berharap penulisan ini dapat bermanfaat bagi kita semua dan bagi penulis pribadi khususnya, serta dapat digunakan sebagaimana mestinya.

Depok, April 2020

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
LEMBAR ORIGINALITAS DAN PUBLIKASI	ii
LEMBAR PENGESAHAN	iii
ABSTRAKSI	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	ix
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
BAB I : PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Batasan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Metode Penelitian	2
1.5 Sistematika Penulisan	3
BAB II : TINJAUAN PUSTAKA	4
2.1 Teorema Penaksiran Universal	4
2.2 Jaringan Saraf Tiruan	4
2.3 Residual Network	5
2.4 Fungsi Aktivasi	5
2.5 Estimasi Pose Dua Dimensi	5
2.6 Estimasi Pose Tiga Dimensi	5
BAB III : PENDEKATAN	6
3.1 Motivasi	6
3.2 Framework Riset	6
3.3 Pendekatan	6

BAB IV : HASIL DAN ANALISIS	7
4.1 Persiapan pengujian	7
4.2 Pelaksanaan Pengujian	7
4.3 Hasil dan Diskusi	7
BAB V : PENUTUP	8
5.1 Kesimpulan	8
5.2 Saran	8
DAFTAR PUSTAKA	9
LAMPIRAN	L1

DAFTAR TABEL

2.1	Sebuah tabel	5
-----	------------------------	---

DAFTAR GAMBAR

2.1	Ilustrasi Sebuah Sel Saraf	5
3.1	Agreement Results on Case 1	6

DAFTAR LAMPIRAN

Lampiran 1 Listing Program	10
----------------------------------	----

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pemanfaatan teknologi yang terkomputerisasi oleh manusia selalu meninggalkan rekam jejak yang tersimpan dalam bentuk data digital. Data digital yang umumnya dimanfaatkan oleh manusia meliputi teks, citra audio, citra visual, dan citra audio visual yang dimampatkan kedalam suatu media penyimpanan. Pengaksesan data ini tergolong mudah karena data dikonstruksi dengan struktur dan format yang sama sehingga dapat diakses dimana saja. Jumlah rekam jejak digital yang tersedia semakin bertambah seiring dengan mudahnya akses terhadap teknologi secara skala besar.

Rekam jejak digital yang terkumpul bersifat laten yang berarti data yang tersedia tidak memiliki semantik. Data ini hanya akan berguna apabila ada suatu aturan yang ditentukan oleh pengguna. Proses pembuatan citra visual terbentuk dari hasil perekaman kamera dengan mengubah gelombang elektromagnetik yang dipantulkan oleh benda dari posisi tertentu menuju lensa kamera menjadi tiga lapis deretan angka numerik yang mewakili warna merah, hijau, dan biru. Citra visual yang tercipta hanya memiliki informasi numerik mengenai warna dalam bentuk gambar, sedangkan informasi posisi benda saat perekaman sudah hilang. Otak manusia dapat mengartikan kondisi suatu benda dalam suatu gambar tanpa kita ketahui cara kerjanya secara pasti. Hal yang sama dapat dilakukan oleh komputer dengan membuat pemetaan dari suatu gambar ke suatu kondisi yang diinginkan.

Permasalahan yang ingin diselesaikan adalah mengembalikan informasi posisi pose tubuh manusia yang telah hilang dengan membuat pemetaan dari gambar yang ditangkap menggunakan kamera monokuler ke koordinat setiap titik kunci dari pose tersebut. Permodelan pembelajaran dalam atau *deep learning* dapat pemetaan suatu domain ke domain lainnya secara otomatis menggunakan pelatihan jaringan saraf tiruan dalam atau *deep neural network*. Fungsi pemetaan dapat dihasilkan dengan melatih jaringan saraf tiruan dengan jumlah data yang besar.

1.2 Batasan Masalah

Penelitian ini menganggap setiap pose dua dimensi maupun pose tiga dimensi berada dalam koordinat lokal. Setiap pose ditransformasi ke dalam observasi kamera dengan titik kunci pinggang sebagai posisi tengah. Hal ini dilakukan karena pemetaan hanya menggunakan grafik datar tanpa informasi kedalaman titik kunci sehingga dapat menghindari masalah kedalaman yang ambigu.

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk membuat sebuah aplikasi yang dapat membaca titik kunci dari sebuah citra visual datar dan melakukan transformasi ke pose lokal dua dimensi sehingga dapat dipetakan oleh jaringan saraf tiruan yang dimodelkan ke bentuk pose lokal tiga dimensi. Pose hasil juga divisualisasikan secara interaktif sehingga dapat dipergunakan untuk kepentingan yang sesuai.

1.4 Metode Penelitian

Penelitian dibagi menjadi tiga tahap besar terurut yang terdiri dari *data preprocessing*, pelatihan model jaringan saraf tiruan, dan visualisasi. Tahap pertama dan tahap ketiga tidak melibatkan pembelajaran mesin. Masalah utama dari penelitian ini berada pada tahap kedua tentang pelatihan jaringan saraf tiruan untuk melakukan pemetaan pose dua dimensi ke pose tiga dimensi.

Dataset yang digunakan dalam penelitian ini adalah *Human3.6M* yang berisi 3,6 juta pose unik yang dilakukan oleh sebelas aktor profesional dan direkam menggunakan empat sudut kamera yang berbeda beserta dengan koordinat setiap titik kunci dari hasil penangkapan alat *motion capture* [4].

Pose dua dimensi dan tiga dimensi merupakan variabel yang relevan untuk masalah. Pada tahap *data preprocessing* akan dilakukan ekstraksi variabel ini kedalam bentuk numerik sehingga mudah untuk digunakan saat melakukan pelatihan jaringan saraf tiruan. Tahap selanjutnya akan dilakukan pembuatan, permodelan, pelatihan, dan evaluasi jaringan saraf tiruan dalam untuk pemetaan

titik kunci yang kemudian dilanjutkan dengan percobaan model dengan video. Tahap terakhir menampilkan visualisasi gambar, pose dua dimensi, dan pose tiga dimensi.

Perangkat keras yang digunakan dalam penelitian ini adalah satu unit laptop dengan spesifikasi:

- CPU Intel Core I7 7700HQ
- Memori 24 GB DDR4
- GPU NVIDIA GTX 1060 6GB
- SSD NVME SAMSUNG 120 GB
- HDD SATA 1 TB

1.5 Sistematika Penulisan

Adapun sistematika penulisan yang digunakan dalam penulisan ini adalah sebagai berikut:

PENDAHULUAN, mengemukakan latar belakang masalah, batasan masalah, tujuan penelitian, metode penelitian, dan sistematika penulisan.

TINJAUAN PUSTAKA, menjelaskan kumpulan teori yang digunakan dalam mendukung proses penyelesaian program.

PENDEKATAN, mengemukakan langkah-langkah yang dicapai untuk membuat program.

HASIL DAN ANALISIS, menghubungkan hasil yang didapatkan dengan teori-teori yang dibahas pada bab 2.

PENUTUP, mengulas lebih lanjut mengenai kesimpulan yang dapat ditarik dari hasil disertai dengan saran yang dapat menyempurnakan penelitian selanjutnya.

BAB II

TINJAUAN PUSTAKA

2.1 Teorema Penaksiran Universal

Teorema penaksiran universal atau *universal approximation theorem* menyatakan bahwa sebuah model jaringan *feed-forward* dapat membentuk fungsi apapun secara subjektif. Sebuah model jaringan saraf tiruan dibentuk dari serangkaian lapisan yang didalamnya terdapat deretan sel saraf atau *neuron* dengan kuantitas tertentu. Semakin panjang rangkaian lapisan yang tersedia, maka semakin banyak saraf yang tersedia sehingga dapat memetakan fungsi yang sulit. Model jaringan yang memiliki banyak saraf dapat mempelajari pola-pola yang ada dari satu domain ke domain lainnya [2].

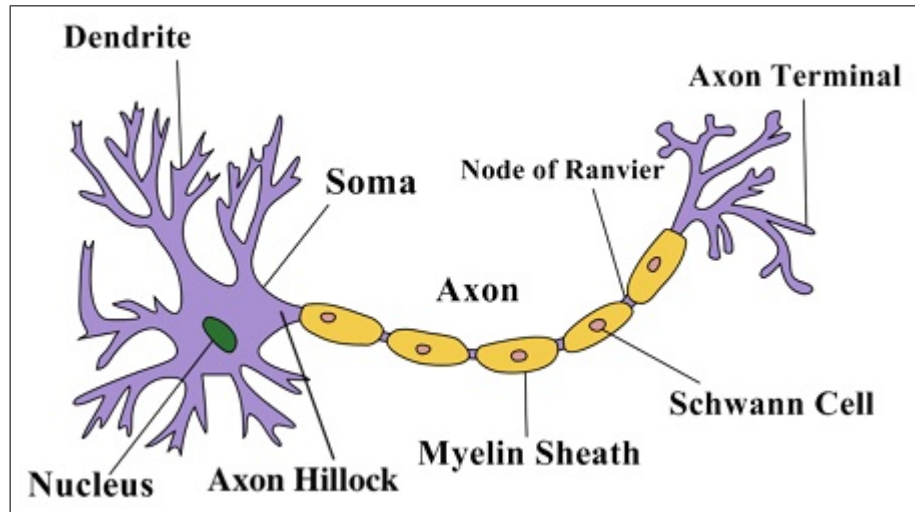
Teorema penaksiran universal memiliki dua sifat yang dikategorikan berdasarkan pemanfaatannya dalam melakukan pembelajaran mesin. Sifat pertama adalah suatu model jaringan saraf tiruan dapat memperkirakan suatu fungsi dengan batasan-batasan tertentu sesuai dengan fungsi aktivasi pada lapisan terakhir. Sifat kedua adalah sebuah fungsi kontinu dengan jumlah variabel sembarang dapat ditiru oleh sebuah jaringan saraf tiruan dengan jumlah yang sembarang [5].

2.2 Jaringan Saraf Tiruan

Penelitian mengenai otak manusia sudah dimulai sejak ribuan tahun yang lalu. Otak manusia terdiri dari kumpulan sel saraf yang saling terkoneksi satu sama lain. Sebuah sel saraf adalah sel yang dapat memproses dan mengantarkan informasi apabila dirangsang dengan tegangan elektrokimia. Sel-sel saraf tidak pernah memperbanyak dirinya dan tidak digantikan apabila ada yang rusak. Jumlah sel saraf yang terdapat dalam otak manusia diperkirakan sebanyak satu miliar. Setiap sel saraf diperkirakan berkoneksi dengan sepuluh ribu sel saraf lainnya melalui sinapsis yang berarti otak manusia beroperasi seperti prosesor dengan kecepatan satu triliun bit per detik [3].

Bentuk sel saraf sangat bervariasi dengan berbagai ukuran, bentuk, dan

sifat elektrokimianya. Sebuah sel saraf memiliki badan yang terdiri dari beberapa struktur penting seperti pada gambar 2.1.



Gambar 2.1: Ilustrasi Sebuah Sel Saraf

Sumber: Sumber yang Berbahaya

Bentuk sel saraf sangat bervariasi dengan berbagai ukuran, bentuk, dan sifat elektrokimianya. Sebuah sel saraf memiliki badan yang terdiri dari beberapa struktur penting seperti pada gambar

2.3 Residual Network

Tabel 2.1: Sebuah tabel

1	2	3
4	5	6
7	8	9

Sumber: Bego Lu

2.4 Fungsi Aktivasi

2.5 Estimasi Pose Dua Dimensi

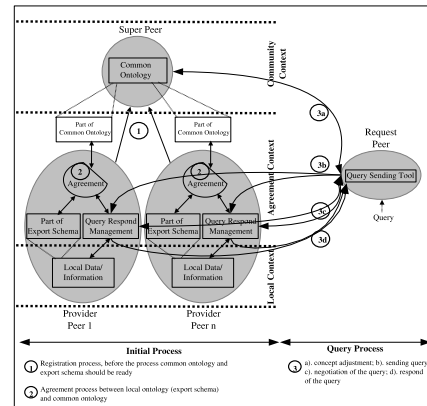
citasi [1]

2.6 Estimasi Pose Tiga Dimensi

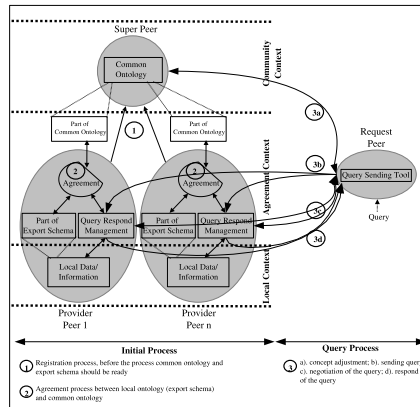
citasi [6]

Information Interoperability & Decision Support Domain Application			
SEMANTIC INTEROP	QUERY PROCESSING	GIS INTEROP	P2P
<ul style="list-style-type: none"> Heterogeneous semantic Semantic description of sources 	<ul style="list-style-type: none"> Query rewriting Query plan and optimization Merge respond 	<ul style="list-style-type: none"> More complex and heterogeneous Spatial ontology 	<ul style="list-style-type: none"> Model Architecture Characteristics
SEMANTIC INTEROP	QUERY PROCESSING	GIS INTEROP	P2P

(a) Agreement by using Algorithm



(b) Agreement by using Algorithm and User feedback



(c) Manual mapping by Cruz et al

Gambar 3.1: Agreement Results on Case 1

BAB III

PENDEKATAN

3.1 Motivasi

Motivasi dari Metodologi

3.2 Framework Riset

Isi tentang framework dari riset

3.3 Pendekatan

BAB IV

HASIL DAN ANALISIS

4.1 Persiapan pengujian

Berisi langkah2 untuk persiapan pengujian, bisa secara pembuktian secara teoritis, empiris, simulasi, dll.

4.2 Pelaksanaan Pengujian

Berisi tentang langkah2 pelaksanaan

4.3 Hasil dan Diskusi

Berisi tentang hasil2 pengujian, ulasan diskusi dari penghasilan dan memberikan penekanan hal yang penting dari pengujian.

BAB V

PENUTUP

5.1 Kesimpulan

Berisi ringkasan dari metodologi dan kesimpulan penting dari hasil evaluasi.

5.2 Saran

Berisi saran-saran untuk pengembangan riset ini ke langkah ke depan.

DAFTAR PUSTAKA

- [1] Z. Cao, G. Hidalgo Martinez, T. Simon, S. Wei, and Y. A. Sheikh. Openpose: Realtime multi-person 2d pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.
- [2] N. J. Guliyev and V. E. Ismailov. A single hidden layer feedforward network with only one neuron in the hidden layer can approximate any univariate function. *arXiv e-prints*, page arXiv:1601.00013, Dec. 2015.
- [3] S. Herculano-Houzel. The human brain in numbers: a linearly scaled-up primate brain. *Frontiers in Human Neuroscience*, 3:31, 2009.
- [4] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu. Human3.6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7):1325–1339, jul 2014.
- [5] A. Kratsios. The Universal Approximation Property: Characterizations, Existence, and a Canonical Topology for Deep-Learning. *arXiv e-prints*, page arXiv:1910.03344, Oct. 2019.
- [6] J. Martinez, R. Hossain, J. Romero, and J. J. Little. A simple yet effective baseline for 3d human pose estimation. In *ICCV*, 2017.
- [7] J. Zhang. Basic Neural Units of the Brain: Neurons, Synapses and Action Potential. *arXiv e-prints*, page arXiv:1906.01703, May 2019.

LAMPIRAN

Bisa diketik sesuai kebutuhan