PENGENALAN AKSI PADA VIDEO MENGGUNAKAN METODE TRANSFORMER

SKRIPSI

Diajukan untuk Memenuhi Bagian dari
Syarat Memperoleh Gelar Sarjana Komputer
Pada Departemen Pendidikan Ilmu Komputer
Program Studi Ilmu Komputer

oleh
Rantty Gantini
NIM 1701266

PROGRAM STUDI ILMU KOMPUTER
DEPARTEMEN PENDIDIKAN ILMU KOMPUTER
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PENDIDIKAN INDONESIA
2022

PENGENALAN AKSI PADA VIDEO MENGGUNAKAN METODE TRANSFORMER

Oleh

Rantty Gantini

Sebuah skripsi yang diajukan untuk memenuhi salah satu syarat memperoleh gelar Sarjana pada Fakultas Pendidikan Matematika dan Ilmu Pengetahuan Alam

© Rantty Gantini

Universitas Pendidikan Indonesia

Juli 2022

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

RANTTY GANTINI

PENGENALAN AKSI PADA VIDEO MENGGUNAKAN METODE TRANSFORMER

disetujui dan disahkan oleh pembimbing:

Pembimbing I

Yaya Wihardi, M. Kom.

NIP. 198903252015041001

Pembimbing II

Enjun Junaeti, M.Si.

NIP. 198512202012122002

Mengetahui,

Ketua Program Studi Ilmu Komputer

Dr. Rani Megasari, M.T.

198705242014042002

PERNYATAAN

Dengan ini saya menyatakan bahwa skripsi dengan judul “Pengenalan Aksi pada Video Menggunakan Metode Transformer” ini beserta seluruh isinya adalah benar-benar karya saya sendiri. Saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika ilmu yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini, saya siap menanggung risiko/sanksi apabila di kemudian hari ditemukan adanya pelanggaran etika keilmuan atau ada klaim dari pihak lain terhadap keaslian karya saya ini.

Bandung, Juli 2022

Yang membuat pernyataan,

Rantty Gantini

KATA PENGANTAR

Puji serta syukur ke hadirat Allah *subhanahuwata'ala* yang telah melimpahkan berkah, rahmat, dan karunia-Nya sehingga penyusunan skripsi yang berjudul "Pengenalan Aksi pada Video Menggunakan Metode Transformer" ini dapat diselesaikan sesuai dengan rencana penelitian.

Tujuan dari penyusunan skripsi ini adalah sebagai salah satu syarat untuk memenuhi syarat untuk memperoleh gelar sarjana komputer pada jenjang studi S1 Program Studi Ilmu Komputer di Universitas Pendidikan Indonesia.

Dalam penyusunan skripsi ini, penulis telah berupaya semaksimal mungkin dalam menyajikan hasil penelitian yang dituangkan dalam bentuk laporan sehingga dapat dibaca oleh pembaca umum. Namun, penulis menyadari bahwa masih banyak kekurangan serta keterbatasan dalam penyusunan tulisan ini. Oleh karena itu, penulis sangat terbuka dan menerima atas kritikan dan saran terhadap tulisan ini agar segala kekurangan dan kualitas yang terdapat pada tulisan ini dapat diperbaiki di tulisan berikutnya.

Bandung, Juli 2022
Penulis,

Rantty Gantini

UCAPAN TERIMA KASIH

Alhamdulillahirabbil'alamin, puji dan syukur kehadirat Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya sehingga penulis diberikan kelancaran dalam menyelesaikan penulisan skripsi ini. Dalam proses menyelesaikan penelitian dan penyusunan skripsi ini, penulis banyak mendapat bimbingan, dorongan, serta bantuan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis mengucapkan terima kasih dengan setulus hati serta penghargaan yang setinggi-tingginya, kepada:

1. Kedua orang tua, dan kakak-kakak penulis yang selalu memberikan doa terbaiknya, juga semangat dan dorongan kepada penulis dalam menjalankan perkuliahan dan menyelesaikan skripsi ini.
2. Bapak Yaya Wihardi, M.Kom., sebagai pembimbing I yang selalu sabar membimbing anak didiknya dengan mengadakan bimbingan rutin setiap pekan. Tidak hanya sebagai pembimbing, namun juga sekaligus mentor terbaik yang tidak hanya memberikan ilmu terbaiknya, namun juga selalu memberikan semangat dan motivasi kepada penulis untuk melanjutkan skripsi yang sempat tertunda ini. Kesempatan yang sangat berharga bisa menjadi mahasiswa bimbingan dari beliau, jasa beliau akan selalu membekas sampai hari nanti.
3. Ibu Enjun Junaeti, M.Si., selaku dosen pembimbing II yang dengan dengan sabarnya memperbaiki segala bentuk penulisan yang telah penulis buat secara detil, terperinci, dan hati-hati.
4. Ibu Dr. Rani Megasari, M.T., selaku ketua Program Studi Ilmu Komputer yang selalu memberikan semangat juga selalu meluangkan waktunya untuk memberikan atensi dan fasilitas khusus kepada mahasiswa yang sedang mengerjakan skripsi, salah satunya dengan mengadakan kegiatan residensi.
5. Bapak Harsa Wara Prabawa, S.Si., M.Pd., selaku dosen Pembimbing Akademik yang dari awal perkuliahan selalu menjadi dosen wali terbaik dengan arahan dan motivasinya, sampai akhir perkuliahan selalu intensif menanyakan capaian kemajuan pengerjaan skripsi mahasiswa akademiknya.

6. Ibu Rosa Ariani Sukamto, S.T., M.T. yang setiap perkuliahannya selalu menyisipkan nilai-nilai kehidupan kepada mahasiswanya untuk tidak perlu menjadi orang lain, selalu jujur dalam berlaku, selalu berjuang dan tidak malu untuk belajar.
7. Bapak Eddy Prasetyo Nugroho, M.T., selaku demisioner ketua Program Studi Ilmu Komputer yang menjadi salah satu dosen yang cukup intensif dalam berkomunikasi dengan penulis ketika mengadakan acara himpunan dan ketika persiapan lomba.
8. Bapak Dr. Lala Septem Riza, M.T., selaku ketua demisioner Departemen Pendidikan Ilmu Komputer yang telah memberikan kesempatan kepada penulis untuk menjadi salah satu tim web dan media sosial untuk Departemen Pendidikan Ilmu Komputer.
9. Seluruh dosen yang tidak bisa penulis tuliskan satu persatu yang telah membimbing dan memberikan ilmu yang bermanfaat pada penulis semasa perkuliahan.
10. Muhammad Rizal Purnomo Putra Perdana sebagai salah satu supporter terbaik yang penulis miliki, yang selalu sabar untuk menemani, memberikan semangat, dorongan, dan bantuan yang begitu banyak dalam hal apapun selama ini.
11. Nur ‘Aisyah Nadiyah dan Fenny Feronika Maharani sebagai sahabat terbaik penulis selama perkuliahan yang selalu memberikan semangat dan bantuan selama proses perkuliahan dan meluangkan waktunya untuk membantu proses penelitian skripsi ini hingga selesai.
12. Teh Yola Nanda Sekar Prima dan Kang Rahman Abdul Rozak sebagai sahabat dan kakak tingkat penulis yang telah memberikan pengalaman sangat berharga untuk penulis dalam mengikuti perlombaan, dan selalu memberikan semangat kepada penulis untuk menyelesaikan skripsi ini.
13. Muhamad Luthfi Naufal sebagai teman seperjuangan dalam menyelesaikan skripsi, Witri Sulanjar dan Ade Prianto sebagai sahabat sekaligus teman satu konsentrasi dalam topik AI yang selalu menemani dan berbagi pengalaman dalam penelitian.

14. Teman-teman satu kelas di Nehalem 2017, dan seluruh teman-teman Departemen Pendidikan Ilmu Komputer angkatan 2017 yang telah bersama-sama melewati banyak rintangan hingga sampai pada titik terbaik yang bisa masing-masing tempuh.
15. Kakak tingkat dan adik tingkat departemen pendidikan ilmu komputer yang telah banyak memberikan penulis pelajaran dan jadi batu loncatan penulis untuk terus belajar.
16. Badan Eksekutif Mahasiswa terkhusus Divisi Komunikasi Teknologi dan Informasi yang telah banyak membuat penulis belajar banyak mengenai organisasi di masa kampus dan membuat penulis terus ingin belajar mengenai pengembangan diri dan manajemen organisasi.
17. Serta seluruh pihak lainnya yang tidak dapat penulis sebutkan satu persatu yang juga telah membantu dan memberi semangat dalam perkuliahan dan proses penyusunan skripsi ini hingga selesai.

Akhir kata, semoga skripsi ini dapat memberi manfaat dan menambah ilmu bagi pembaca. Tidak ada kata yang dapat menggambarkan rasa terima kasih penulis atas semua dukungan yang telah diberikan, semoga senantiasa diberikan kebaikan, kebahagian, dan kesehatan oleh Allah SWT. *Aamiin.*

PENGENALAN AKSI PADA VIDEO MENGGUNAKAN METODE TRANSFORMER

Oleh

Rantty Gantini – rantigantini11@upi.edu

1701266

ABSTRAK

Manusia dapat dengan mudah mengenali dan mengidentifikasi suatu tindakan atau aksi seseorang dalam sebuah video, namun tidak demikian dengan mesin. Hal ini sulit dilakukan karena beberapa hal, seperti adanya variasi dalam melakukan suatu aksi, objeknya dapat berubah, atau sudut pandang yang berbeda. Teknik *action recognition* dapat menganalisis isi video untuk dapat menentukan tindakan apa saja yang ada pada video tersebut. *Action recognition* ini dapat dimanfaatkan di berbagai bidang seperti *security surveillance*, *autonomous driving*, *robotics*, *early warning*, *video retrieval*, dan *healthcare*. Dari permasalahan tersebut, penelitian ini bertujuan untuk membangun model sistem pengenalan aksi pada video menggunakan metode Transformer. Model ini dibangun menggunakan set data publik UTD-MHAD versi video RGB yang berisikan 861 video aksi sederhana manusia. *Recognizer* pada penelitian ini menggunakan metode Transformer yang dapat menangani permasalahan pada data sekuensial. Akurasi tertinggi yang berhasil didapatkan pada penelitian ini adalah 83,78%.

Kata Kunci: Computer Vision, Deep Learning, Action Recognition, Transformer.

PENGENALAN AKSI PADA VIDEO MENGGUNAKAN METODE TRANSFORMER

Arranged by

Rantty Gantini – rantigantini11@upi.edu

1701266

ABSTRACT

Humans can easily recognize and identify an action or someone's actions in a video, but not so with machines. This is difficult to do for several reasons, such as the variation in performing an action, the object may change, or a different point of view. The action recognition technique can analyze the content of the video to determine what actions are in the video. This action recognition can be used in various fields such as security surveillance, autonomous driving, robotics, early warning, video retrieval, and healthcare. From these problems, this study aims to build a model of an action recognition system on video using the Transformer method. This model was built using the RGB video version of the UTD-MHAD public dataset containing 861 simple human action videos. The recognizer in this study uses the Transformer method which can handle problems in sequential data. The highest accuracy that was obtained in this study was 83.78%.

Keyword: Computer Vision, Deep Learning, Action Recognition, Transformer.

DAFTAR ISI

HALAMAN PENGESAHAN i
HALAMAN PERNYATAAN ii
KATA PENGANTAR iii
UCAPAN TERIMA KASIH iv
ABSTRAK vii
ABSTRACT viii
DAFTAR ISI ix
DAFTAR GAMBAR xi
DAFTAR TABEL xi
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 3
1.3 Tujuan Penelitian 3
1.4 Batasan Masalah 3
1.5 Manfaat Penelitian 4
1.6 Sistematika Penulisan 4
BAB II KAJIAN PUSTAKA 6
2.1 Peta Literatur 6
2.2 Penelitian Terkait 6
2.3 *Computer Vision* 9
2.4 Pengolahan Citra Digital 11
2.5 *Action Recognition* 13
2.6 Jaringan Saraf Tiruan 15
2.7 *Deep Learning* 16
2.8 Metode *Recurrent Neural Network* (RNN) 17
2.9 Metode Transformer 18
2.9.1 Encoder 20
2.9.2 Decoder 21
2.9.3 Multi-Head Attention 21

2.9.4 Jaringan Feed-Forward 22
BAB III METODOLOGI PENELITIAN 23
3.1 Desain Penelitian 23
3.2 Data 24
3.3 Eksperimen 27
3.4 Evaluasi dan Analisis 31
3.5 Lingkungan Komputasi 31
BAB IV HASIL DAN PEMBAHASAN 32
4.1 Implementasi 32
4.1.1 Preprocessing 32
4.1.2 Feature Extraction 35
4.1.3 Split *Data* Training *dan* Testing 35
4.1.4 Construct Model 35
4.1.5 Training 36
4.1.6 Hasil Eksperimen Proses Training 36
4.2 Evaluasi dan Analisis 38
BAB V KESIMPULAN DAN REKOMENDASI 42
5.1 Kesimpulan 42
5.2 Rekomendasi 42
DAFTAR PUSTAKA 44

DAFTAR GAMBAR

Gambar 2.1 Peta Literatur Penelitian 6
Gambar 2.2 *Optical Character Recognition* (Majumder et al., 2019) 10
Gambar 2.3 Contoh *Surveillance System* (Handojo et al., 2006) 10
Gambar 2.4 Koordinat Citra Digital (Ratna et al., 2016) 11
Gambar 2.5 Contoh *Image Processing* pada Citra (Szeliski, 2010) 12
Gambar 2.6 Contoh *Action Recognition* (John, 2019) 13
Gambar 2.7 Arsitektur Jaringan *Backpropagation* (Ayu, 2019) 16
Gambar 2.8 Arsitektur Model Metode Transformer (Vaswani et al., 2017) 19
Gambar 2.9 *Residual Connection* dan *Layer Normalization* pada sublapisan 20
Gambar 2.10 Arsitektur *Multi-Head Attention* 22
Gambar 3.1 Desain Penelitian 23
Gambar 3.2 Contoh Set Data UTD-MHAD 26
Gambar 3.3 Contoh Set Data UTD-MHAD yang Terunduh 26
Gambar 3.4 Contoh Penamaan File pada Video 27
Gambar 3.5 Desain Eksperimen 28
Gambar 4.1 Grafik Jumlah Frame pada Video 33
Gambar 4.2 Proses *Cropping* 34
Gambar 4.3 Proses *Resizing* 34
Gambar 4.4 *Confusion Matrix* 39

DAFTAR TABEL

Tabel 3.1 Nama Aksi dalam Video 25
Tabel 4.1 *Tuning Hyperparameter* 36
Tabel 4.2 Akurasi untuk Setiap Model 36
Tabel 4.3 Hasil Prediksi Label 37