

**KLASIFIKASI MULTI-LABEL PADA GENRE NOVEL
BERDASARKAN COVER MENGGUNAKAN METODE
CONVOLUTIONAL NEURAL NETWORK**

*Diajukan Sebagai Syarat Untuk Menyelesaikan
Pendidikan Program Strata-1 Pada
Jurusan Teknik Informatika*



Oleh:

Yasmin Azzahrah Lubis

NIM: 09021181722017

**Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2022**

LEMBAR PENGESAHAN SKRIPSI

KLASIFIKASI MULTI-LABEL PADA GENRE NOVEL BERDASARKAN *COVER* MENGGUNAKAN METODE *CONVOLUTIONAL NEURAL NETWORK*

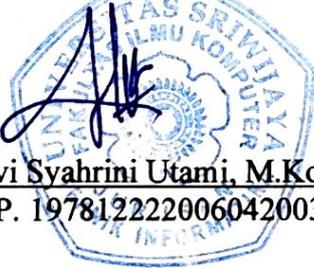
Oleh:

Yasmin Azzahrah Lubis

NIM: 09021181722017

Palembang, 24 Januari 2022

Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003

Pembimbing,

A handwritten signature in black ink, consisting of a stylized 'J' followed by a long horizontal stroke and a vertical stroke.

Julian Supardi, M.T.
NIP. 197207102010121001

TANDA LULUS UJIAN SIDANG TUGAS AKHIR

Pada hari Senin tanggal 3 Januari 2022 telah dilaksanakan ujian sidang tugas akhir oleh Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Nama : Yasmin Azzahrah Lubis
NIM : 09021181722017
Judul : Klasifikasi Multi Label Pada Genre Novel Berdasarkan Cover Menggunakan Metode *Convolutional Neural Network*

1. Pembimbing

Julian Supardi, M.T.
NIP. 197207102010121001



.....

2. Penguji I

Samsuryadi, M.Kom., Ph.D.
NIP. 197102041997021003



.....

3. Penguji II

Muhammad Naufal Rachmatullah, M.T.
NIP. 1671060112920006



.....

Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M.Kom.
NIP. 197812222006042003

HALAMAN PERNYATAAN

Yang bertanda tangan dibawah ini:

Nama : Yasmin Azzahrah Lubis
NIM : 09021181722017
Program Studi : Teknik Informatika
Judul : Klasifikasi Multi Label Pada Genre Novel Berdasarkan
*Cover Menggunakan Metode Convolutional Neural
Network*

Hasil pengecekan *software ithenticate/turnitin* : 4%

Menyatakan bahwa laporan tugas akhir saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan tugas akhir ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapa pun.

Palembang, 24 Januari 2022



Yasmin Azzahrah Lubis
NIM. 09021181722017

MOTTO DAN PERSEMBAHAN

Motto

“Sesungguhnya Aku sesuai dengan persangkaan hamba-Ku kepada-Ku.”

[HR. Al-Bukhari & Muslim]

“One must be sane to think clearly, but one can think deeply and be quite insane.”

[Nikola Tesla]

“If you see it in your mind, you’re going to hold it in your hand.”

[Bob Proctor]

“The thoughts you think and the way you feel are at the center of what you attract”

[Abraham Hicks]

Kupersembahkan karya tulis ini kepada:

- Allah SWT
- Kedua Orang Tuaku
- Keluarga Besarku
- Dosen Pembimbingku
- Almamaterku

MULTI LABEL CLASSIFICATION OF FICTION BOOK GENRE BASED ON ITS COVER USING CONVOLUTIONAL NEURAL NETWORK

By:
Yasmin Azzahrah Lubis
09021181722017

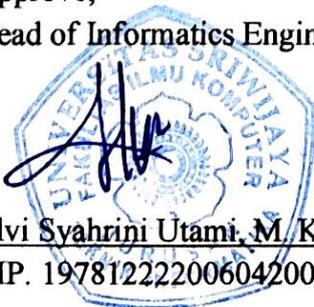
ABSTRACT

Fiction book classification based on its cover become a particularly exciting research topic in recent years. There exists a wide variety of book genres, and the design on its cover as a graphic designs vary in many different ways such as colors, styles, textual information, even for books of the same genre, make it an challenging task to do fiction book genre classification based on its cover. In some related work before, classification system for define genre of fiction book practice still limited with traditional classification, where a fiction book will only be classified into one genre, whereas in real situations a novel can be classified into more than one genre. In addition, another problem lies in the data redundancy in several categories of the dataset. Therefore, in this study will try to solved the two problems, the first by applying multi-label classification practice with a Convolutional Neural Network model in order to classify the fiction book genre information based on the cover and provide the correct genre classification results, then the second we will collect fiction book's cover image as new dataset using web scraping technique, so the data can be used for multi label classification purpose. The Convolutional Neural Network model used in this study is the VGG16 architecture. The results of experiments conducted in this study indicate that the VGG16 model that has been trained using the dataset that has been collected by the author can perform a multi label classification to determine the novel genre in the testing process properly and obtain a fairly good F1 Score.

Key words: *Multi Label Classification, Fiction Book Genre, Cover, Deep Learning, Convolutional Neural Network, VGG16*

Palembang, January 24th 2022

Approve,
Head of Informatics Engineering Department



Alvi Syahrini Utami, M. Kom
NIP. 197812222006042003

Supervisor,

Julian Supardi, M.T
NIP. 197207102010121001

KLASIFIKASI MULTI-LABEL PADA GENRE NOVEL BERDASARKAN COVER MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK

Oleh:

Yasmin Azzahrah Lubis

09021181722017

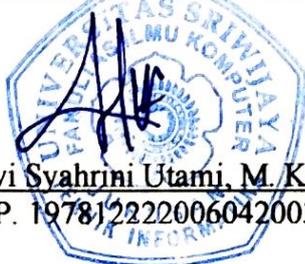
ABSTRAK

Klasifikasi genre novel berbasis sampul menjadi topik penelitian yang sangat menarik dalam beberapa tahun terakhir. Adanya berbagai ragam genre novel yang berbeda, serta desain pada sampul novel yang dapat bervariasi dari segi warna, gaya, dan informasi tekstual, membuat hal ini menjadi tantangan dalam melakukan klasifikasi genre novel berdasarkan sampulnya. Pada penelitian yang dilakukan sebelumnya, sistem klasifikasi genre novel berdasarkan gambar sampulnya masih terbatas dengan penerapan klasifikasi biasa, dimana sebuah novel hanya akan digolongkan ke satu genre saja, sedangkan pada keadaan nyata sebuah novel dapat tergolong kedalam lebih dari satu genre. Selain itu, permasalahan lain juga terletak pada data yang digunakan terdapat redundansi pada beberapa kategori. Oleh karena itu, pada penelitian ini akan dilakukan penyelesaian kedua masalah tersebut, yaitu yang pertama dengan melakukan penerapan klasifikasi multi label menggunakan model *Convolutional Neural Network* agar dapat mengelompokkan genre novel berdasarkan sampulnya dan memberikan hasil klasifikasi genre dengan tepat, kemudian yang kedua dengan mengumpulkan data set sampul novel yang baru menggunakan teknik *web scraping* agar data yang dipakai dapat sesuai dengan kebutuhan sistem klasifikasi multi label yang akan dilakukan. Adapun model *Convolutional Neural Network* yang digunakan adalah arsitektur VGG16. Hasil percobaan yang dilakukan pada penelitian ini menunjukkan bahwa model VGG16 yang telah dilatih menggunakan data set yang telah dikumpulkan penulis dapat melakukan klasifikasi multi label untuk menentukan genre novel pada proses pengujian dengan baik dan memperoleh nilai *F1 Score* yang cukup baik.

Kata Kunci: Klasifikasi Multi Label, Genre Novel, Sampul, *Deep Learning*, *Convolutional Neural Network*, VGG16

Palembang, 24 Januari 2022

Mengetahui,
Ketua Jurusan Teknik Informatika



Alvi Syahrini Utami, M. Kom
NIP. 197812222006042003

Pembimbing,

Julian Supardi, M.T
NIP. 197207102010121001

KATA PENGANTAR

Alhamdulillah Robbil'Alamin, dengan menyebut nama Allah SWT. Saya berterima kasih atas semua rahmat dan karunia-Nya karena berkat petunjuk dan kemudahan dari-Nya sehingga saya dapat menyelesaikan Tugas Akhir saya yang berjudul “Klasifikasi Multi-Label Pada Genre Novel Berdasarkan *Cover* Menggunakan Metode *Convolutional Neural Network*” dengan baik. Tugas Akhir ini disusun sebagai salah satu syarat dalam menyelesaikan pendidikan program Strata-1 Program Studi Teknik Informatika pada Fakultas Ilmu Komputer di Universitas Sriwijaya.

Dalam menyelesaikan Tugas Akhir ini terdapat banyak pihak yang telah memberikan bantuan dan dukungan baik secara langsung maupun tidak langsung. Oleh karena itu, saya ingin menyampaikan rasa terima kasih yang tulus kepada:

1. Kedua orang tua saya, Ir. Akhmad Ferdinand Lubis dan Jeanny Aprianty Ardjan, S.E., dan adik-adikku, yang senantiasa mendoakan dan selalu memberikan dukungan yang luar biasa, serta pengorbanan yang tak terhingga agar saya selalu mendapatkan yang terbaik.
2. Bapak Jaidan Jauhari, S.Pd., M.T. selaku Dekan Fakultas Ilmu Komputer Universitas Sriwijaya.
3. Ibu Alvi Syahrini Utami, M.Kom. selaku Ketua Jurusan Teknik Informatika di Fakultas Ilmu Komputer Universitas Sriwijaya.

4. Bapak Julian Supardi, M.T. selaku dosen pembimbing saya yang telah memberikan arahan dan kemudahan, dan sudah banyak sekali membantu dan membimbing saya dengan sabar dalam penyusunan tugas akhir saya.
5. Bapak Samsuryadi, M.Kom., Ph.D. dan Bapak Muhammad Naufal Rachmatullah, M.T. selaku dosen penguji tugas akhir saya yang sudah membantu memberikan saran dan masukan dalam penyusunan tugas akhir.
6. Mba Wiwin Juliani selaku admin Teknik Informatika Bilingual yang sudah banyak sekali membantu dan memberikan kemudahan kepada saya dalam pengurusan berkas tugas akhir ini.
7. Seluruh Bapak dan Ibu Dosen Program Studi Teknik Informatika beserta para staf Fakultas Ilmu Komputer Universitas Sriwijaya yang telah memberikan ilmunya selama masa kegiatan perkuliahan saya.
8. Teman-teman saya Fitri Amelia, Mega Pertiwi, Nanda Tirana Mahesa, dan Suci Amelia yang sudah menerima saya, senantiasa selalu setia mendengarkan keluh kesah dan rutukan saya, memberikan banyak bantuan, dukungan, semangat, dan selalu ada untuk saya selama masa perkuliahan ini.
9. Yuni Masrona dan Vira Meilinda, teman-teman saya sejak SMA yang masih selalu memberikan saya dukungan dan nasihat, dan mendengarkan keluh kesah saya selama ini.
10. Rusmansyah Putra NH dan M. Aldi Ariqi, teman-teman yang sudah selalu berbaik hati, banyak membantu saya, dan bersedia direpotkan selama masa perkuliahan dan pengerjaan skripsi saya.

11. Rizq Khairi Yazid, M. Hatta Aldino, dan Vincen, teman-teman VGG16 senasib yang merangkap mentor yang sudah banyak sekali saya repotkan dengan berbagai pertanyaan dan memberikan saran dan masukan untuk saya dalam proses pengerjaan skripsi ini.
12. Semua pihak lainnya yang tidak dapat disebutkan satu-persatu yang telah memberikan doa dan dukungan, serta banyak berperan bagi penulis terutama dalam penyelesaian tugas akhir ini.
13. Dan terutama teruntuk diri saya sendiri yang selama ini sudah selalu berusaha, bekerja mengarahkan semua kemampuan, dan tidak menyerah, walaupun ada banyak sekali masalah dan beban yang muncul selama mengerjakan tugas akhir ini. Terima kasih, Yasmin.

Dalam proses penyelesaian tugas akhir ini saya sadar bahwa masih terdapat banyak kekurangan dalam pengetahuan dan ilmu yang terbatas. Saya akhiri dengan segala hormat, semoga tugas akhir ini dapat memberikan manfaat yang baik untuk kita semua khususnya mahasiswa di Jurusan Teknik Informatika Fakultas Ilmu Komputer Universitas Sriwijaya.

Palembang, 24 Januari 2022



Yasmin Azzahrah Lubis

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
LEMBAR PENGESAHAN SKRIPSI.....	ii
TANDA LULUS UJIAN SIDANG TUGAS AKHIR	iii
HALAMAN PERNYATAAN	iv
MOTTO DAN PERSEMBAHAN	v
ABSTRACT	vi
ABSTRAK	vii
KATA PENGANTAR	viii
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xv
DAFTAR GAMBAR	xvii
DAFTAR ALGORITMA.....	xix
DAFTAR LAMPIRAN	xx
BAB I PENDAHULUAN	
1.1 Pendahuluan	I-1
1.2 Latar Belakang	I-1
1.3 Rumusan Masalah	I-3
1.4 Tujuan Penelitian	I-4
1.5 Manfaat Penelitian	I-4
1.6 Batasan Masalah.....	I-5
1.7 Sistematika Penulisan.....	I-5
1.8 Ringkasan.....	I-7
BAB II KAJIAN LITERATUR	
2.1 Pendahuluan	II-1
2.2 Landasan Teori.....	II-1
2.2.1 Klasifikasi Multi Label	II-1

2.2.1.1 Fungsi Aktivasi pada <i>Output Layer</i>	II-2
2.2.1.2 <i>Evaluation Metrics</i>	II-3
2.2.2 <i>Convolutional Neural Network</i>	II-5
2.2.2.1 Konsep Kerja <i>Convolutional Neural Network</i>	II-6
2.2.2.2 Arsitektur <i>Convolutional Neural Network</i>	II-6
2.2.2.3 VGG16.....	II-8
2.2.2.4 Fungsi Aktivasi.....	II-9
2.2.3 Teknologi Pendukung	II-10
2.2.3.1 Tensorflow.....	II-10
2.2.3.2 Keras	II-11
2.3 Penelitian Lain yang Relevan.....	II-12
2.4 Ringkasan.....	II-14

BAB III METODOLOGI PENELITIAN

3.1 Pendahuluan	III-1
3.2 Pengumpulan Data	III-1
3.3.1 Jenis dan Sumber Data.....	III-1
3.3.2 Metode Pengumpulan Data.....	III-2
3.3 Tahapan Penelitian	III-4
3.3.1 Arsitektur Jaringan.....	III-5
3.3.2 Kerangka Kerja	III-6
3.3.3 Kriteria Pengujian	III-8
3.3.4 Format Data Pengujian	III-8
3.3.5 Alat yang Digunakan dalam Pelaksanaan Pengujian.....	III-10
3.3.6 Pengujian Penelitian.....	III-11
3.3.7 Analisis Hasil Pengujian dan Membuat Kesimpulan.....	III-11
3.4 Metode Pengembangan Perangkat Lunak	III-12
3.4.1 Analisa Kebutuhan.....	III-12
3.4.2 Membangun Rancangan <i>Prototype</i>	III-13
3.4.3 Evaluasi <i>Prototype</i>	III-13
3.4.4 Membangun Sistem	III-13

3.4.5 Pengujian Sistem.....	III-14
3.4.6 Evaluasi Sistem.....	III-14
3.4.7 Menggunakan Sistem.....	III-14
3.5 Manajemen Proyek Penelitian.....	III-16

BAB IV PENGEMBANGAN PERANGKAT LUNAK

4.1 Pendahuluan.....	IV-1
4.2 <i>Prototyping</i>	IV-1
4.2.1 Analisa Kebutuhan.....	IV-1
4.2.1.1 Menentukan Data yang Digunakan	IV-1
4.2.1.2 Menentukan <i>Input</i> dan <i>Output</i>	IV-3
4.2.1.3 Kebutuhan Fungsional dan Non Fungsional	IV-3
4.2.1.4 Spesifikasi Kebutuhan Perangkat Pendukung	IV-5
4.2.2 Membangun Rancangan <i>Prototype</i>	IV-6
4.2.2.1 Implementasi Kebutuhan.....	IV-6
4.2.2.2 Rancangan Antarmuka	IV-17
4.2.3 Evaluasi <i>Prototype</i>	IV-18
4.2.4 Membangun Sistem	IV-20
4.2.5 Menguji Sistem	IV-25
4.2.6 Evaluasi Sistem.....	IV-28
4.2.7 Menggunakan Sistem.....	IV-28
4.3 Ringkasan.....	IV-31

BAB V HASIL DAN ANALISIS PENELITIAN

5.1 Pendahuluan.....	V-1
5.2 Hasil Penelitian.....	V-1
5.2.1 Konfigurasi Penelitian	V-1
5.2.1.1 Konfigurasi Proses Pelatihan.....	V-1
5.2.1.2 Konfigurasi Proses Pengujian.....	V-4
5.2.2 Hasil Proses Pelatihan Data Latih.....	V-4
5.2.3 Hasil Proses Pengujian Data Uji	V-8

5.2.4 Hasil Percobaan Perangkat Lunak Menggunakan Data Sampel	V-14
5.3 Analisis Hasil Penelitian	V-19
5.4 Ringkasan.....	V-25

BAB VI KESIMPULAN DAN SARAN

6.1 Pendahuluan.....	VI-1
6.2 Kesimpulan	VI-1
6.3 Saran	VI-2

DAFTAR PUSTAKA	xxi
----------------------	-----

LAMPIRAN.....	xxiv
---------------	------

DAFTAR TABEL

	Halaman
Tabel II-1	Tabel <i>Confusion Matrix</i> II-3
Tabel III-1	Rancangan Tabel Hasil Pelatihan Model Arsitektur VGG16... III-8
Tabel III-2	Rancangan Tabel <i>Confusion Matrix</i> untuk Klasifikasi Multi Label terhadap Data Uji III-9
Tabel III-3	Rancangan Tabel Hasil Pengujian Prediksi Genre dan Nilai Persentase terhadap Data Sampel III-9
Tabel III-4	Rancangan Tabel Perhitungan Metrik Evaluasi dari Proses Pengujian III-10
Tabel III-5	Tabel Pengembangan Perangkat Lunak Menggunakan Metode <i>Prototyping</i> III-15
Tabel III-6	Tabel <i>Work Breakdown Structure</i> (WBS) Penelitian III-16
Tabel IV-1	Tabel Kebutuhan Fungsional IV-4
Tabel IV-2	Tabel Kebutuhan Non Fungsional IV-4
Tabel IV-3	Tabel Spesifikasi Kebutuhan Perangkat Pendukung IV-5
Tabel IV-4	Definisi Aktor IV-9
Tabel IV-5	Definisi <i>Use-Case</i> IV-9
Tabel IV-6	Skenario <i>Use-Case</i> IV-10
Tabel IV-7	Tabel Implementasi Pembangunan Sistem IV-20
Tabel IV-8	Tabel Pengujian Perangkat Lunak Klasifikasi Genre IV-25
Tabel V-1	Tabel Hasil Pelatihan Model Arsitektur VGG16 V-5
Tabel V-2	<i>Confusion Matrix</i> Genre Fantasi <i>Threshold</i> 0,2 V-9
Tabel V-3	<i>Confusion Matrix</i> Genre General <i>Threshold</i> 0,2 V-9
Tabel V-4	<i>Confusion Matrix</i> Genre Horor <i>Threshold</i> 0,2 V-9
Tabel V-5	<i>Confusion Matrix</i> Genre Misteri <i>Threshold</i> 0,2 V-10
Tabel V-6	<i>Confusion Matrix</i> Genre Romansa <i>Threshold</i> 0,2 V-10
Tabel V-7	<i>Confusion Matrix</i> Genre Fantasi <i>Threshold</i> 0,3 V-10
Tabel V-8	<i>Confusion Matrix</i> Genre General <i>Threshold</i> 0,3 V-10
Tabel V-9	<i>Confusion Matrix</i> Genre Horor <i>Threshold</i> 0,3 V-11

Tabel V-10	<i>Confusion Matrix</i> Genre Misteri <i>Threshold</i> 0,3.....	V-11
Tabel V-11	<i>Confusion Matrix</i> Genre Romansa <i>Threshold</i> 0,3	V-11
Tabel V-12	<i>Confusion Matrix</i> Genre Fantasi <i>Threshold</i> 0,5.....	V-11
Tabel V-13	<i>Confusion Matrix</i> Genre General <i>Threshold</i> 0,5	V-12
Tabel V-14	<i>Confusion Matrix</i> Genre Horor <i>Threshold</i> 0,5.....	V-12
Tabel V-15	<i>Confusion Matrix</i> Genre Misteri <i>Threshold</i> 0,5.....	V-12
Tabel V-16	<i>Confusion Matrix</i> Genre Romansa <i>Threshold</i> 0,5	V-12
Tabel V-17	Hasil Perhitungan Metrik Evaluasi <i>Precision</i> , <i>Recall</i> , dan <i>F1 Score</i> di <i>Threshold</i> 0,2	V-13
Tabel V-18	Hasil Perhitungan Metrik Evaluasi <i>Precision</i> , <i>Recall</i> , dan <i>F1 Score</i> di <i>Threshold</i> 0,3	V-13
Tabel V-19	Hasil Perhitungan Metrik Evaluasi <i>Precision</i> , <i>Recall</i> , dan <i>F1 Score</i> di <i>Threshold</i> 0,5	V-13
Tabel V-20	Hasil Percobaan Klasifikasi Multi Label Menggunakan Data Sampel.....	V-15

DAFTAR GAMBAR

	Halaman
Gambar II-1	<i>Output Node</i> pada Klasifikasi Multi Label II-2
Gambar II-2	Proses Konvolusi..... II-7
Gambar II-3	Proses <i>Max Pooling</i> II-7
Gambar II-4	Model Arsitektur <i>Convolutional Neural Network VGG16</i> .. II-8
Gambar II-5	Struktur Keras dan Tensorflow II-11
Gambar III-1	Tahapan Penelitian III-4
Gambar III-2	Arsitektur Jaringan III-5
Gambar III-3	Kerangka Kerja Penelitian III-6
Gambar III-4	Diagram Metode <i>SDLC Prototyping</i> III-12
Gambar IV-1	Gambar Sampul Novel pada Data Set per Genre..... IV-2
Gambar IV-2	Diagram <i>Flowchart</i> dari Perangkat Lunak..... IV-7
Gambar IV-3	Diagram <i>Use-Case</i> dari Perangkat Lunak..... IV-8
Gambar IV-4	Diagram Aktivasi Klasifikasi Genre dari Sampul Novel..... IV-12
Gambar IV-5	Diagram <i>Data Flow</i> Level 0 IV-13
Gambar IV-6	Diagram <i>Data Flow</i> Level 1 IV-13
Gambar IV-7	Diagram <i>Data Flow</i> Level 2 IV-14
Gambar IV-8	Rancangan Antarmuka <i>Frame</i> Utama..... IV-17
Gambar IV-9	Rancangan Antarmuka <i>Frame</i> Hasil Prediksi..... IV-18
Gambar IV-10	Implementasi Antarmuka <i>Frame</i> Utama..... IV-23
Gambar IV-11	Implementasi Antarmuka <i>Frame</i> Hasil Prediksi..... IV-24
Gambar IV-12	Membuka Terminal Sistem Operasi Windows IV-29
Gambar IV-13	Tampilan <i>Frame</i> Utama IV-29
Gambar IV-14	Tampilan <i>File Dialog</i> dari Perangkat Ketika Pengguna Menekan Tombol ' <i>Browse</i> ' IV-30
Gambar IV-15	Tampilan <i>Frame</i> Hasil Prediksi IV-30
Gambar IV-16	Melakukan Percobaan Klasifikasi Memakai Data Lain Dengan Menekan Tombol ' <i>Try Again</i> ' IV-31
Gambar IV-17	<i>Frame</i> Hasil Prediksi Genre Data Sampul Novel Lain IV-31

Gambar V-1	<i>Summary</i> dari Model VGG16 yang Dibangun.....	V-3
Gambar V-2	Plot Nilai Akurasi dari Proses Pelatihan Model VGG16	V-7
Gambar V-3	Plot Nilai Loss dari Proses Pelatihan Model VGG16	V-7

DAFTAR ALGORITMA

	Halaman
Algoritma III-1 Implementasi <i>Web Scraping</i> Data Set Penelitian	III-3
Algoritma V-1 <i>Threshold</i> Mengubah Nilai <i>Probability</i> ke <i>Multi-Hot</i>	V-8

DAFTAR LAMPIRAN

	Halaman
Lampiran 1 Dokumentasi <i>Source Code</i>	xvii

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab ini akan membahas penjelasan umum mengenai keseluruhan penelitian berdasarkan latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, sistematika penulisan serta kesimpulan.

1.2 Latar Belakang

Sistem klasifikasi genre novel berbasis sampul menjadi topik penelitian yang sangat menarik dalam beberapa tahun terakhir (Kundu & Zheng, 2020). Sudah ada beberapa penelitian terdahulu melakukan klasifikasi genre novel berdasarkan sampulnya dengan menggunakan beberapa model *deep learning* yang saat ini semakin berkembang pesat. Seperti pada penelitian yang dilakukan oleh Chiang dan tim (2015), dimana mereka menggunakan model *Convolutional Neural Network* dengan implementasi *transfer learning* untuk mengolah data gambar sampul novel, dan dua buah algoritma *Natural Language Processing*, *word2vec* dan *Stanford NLP Classifier*, untuk mengolah data teks pada sampul novel. Kemudian ada pula penelitian yang dilakukan oleh Buczkowski dan tim (2018) yang melakukan klasifikasi genre novel hanya berdasarkan data gambar sampulnya saja dengan menggunakan dua buah model *Convolutional Neural Network*, yaitu model dengan konsep VGG dan model arsitektur CNN sederhana, dimana keduanya menghasilkan nilai akurasi yang cukup memuaskan.

Namun, pada data set yang digunakan dalam penelitian sebelumnya terdapat banyak data yang redundansi pada beberapa kategori genre. Hal ini disebabkan karena pada keadaan nyata sebuah novel dapat tergolong kedalam lebih dari satu genre. Sedangkan penerapan pada penelitian-penelitian sebelumnya masih terbatas dengan klasifikasi biasa, dimana sebuah novel hanya akan digolongkan ke satu genre saja. Permasalahan ini dapat diselesaikan dengan melakukan penerapan klasifikasi multi label, dimana nantinya sebuah novel akan dapat dikelompokkan kedalam lebih dari satu genre, tidak hanya terbatas ke satu genre saja seperti pada penelitian sebelumnya. Selain itu, agar dapat melakukan klasifikasi multi label dan menyelesaikan permasalahan data yang redundansi, diperlukan data set gambar sampul novel yang baru yang tidak terdapat duplikat pada tiap kategorinya, agar sesuai dengan kebutuhan penerapan klasifikasi multi label yang akan dilakukan.

Pada beberapa penelitian yang dilakukan sebelumnya, model *Convolutional Neural Network* dapat diimplementasikan dengan baik dalam menangani masalah pengolahan gambar, termasuk sistem klasifikasi genre novel berdasarkan gambarnya. Salah satunya adalah arsitektur VGG16 yang akan digunakan dalam penelitian ini. Arsitektur VGG16 memiliki kedalaman *layer* yang tinggi sehingga kemampuan ekstraksi fiturnya sudah terbukti baik. Selain itu arsitektur VGG16 ini mudah untuk diimplementasikan dan tepat untuk tugas pengolahan gambar. Seperti pada penelitian yang dilakukan oleh Lydia dan Francis (2020), dimana mereka melakukan klasifikasi multi label pada gambar pakaian dan sayur yang diambil dari Google Image dengan menggunakan arsitektur VGG16, karena model VGG16 sudah terbukti memiliki performa yang stabil.

Oleh karena itu, berdasarkan latar belakang dan penjelasan yang telah diuraikan, untuk mengatasi permasalahan yang ada, pada penelitian ini akan dilakukan pengumpulan data set gambar sampul novel yang baru yang akan dikumpulkan menggunakan teknik *web scraping*, dan melakukan penerapan klasifikasi multi label untuk menentukan genre novel berdasarkan gambar sampulnya dengan tepat.

1.3 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan sebelumnya, penerapan klasifikasi genre novel berdasarkan sampulnya yang dilakukan pada penelitian-penelitian terdahulu memiliki tingkat yang memuaskan. Namun, pada data set yang digunakan dalam penelitian sebelumnya terdapat banyak data yang redundansi pada beberapa kategori genre. Hal ini disebabkan karena pada keadaan nyata sebuah novel dapat tergolong kedalam lebih dari satu genre. Sedangkan pada penelitian yang dilakukan sebelumnya masih terbatas dengan penerapan klasifikasi biasa dimana sebuah novel hanya akan digolongkan ke satu genre saja. Oleh karena itu untuk mengatasi permasalahan yang ada, diperlukan penerapan klasifikasi multi label menggunakan arsitektur *Convolutional Neural Network* agar dapat mengelompokkan genre novel berdasarkan sampulnya dan memberikan hasil klasifikasi genre yang tepat. Kemudian juga diperlukan data set gambar sampul novel yang baru agar data yang dipakai dapat sesuai dengan kebutuhan sistem klasifikasi multi label yang akan dilakukan.

1.4 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Mengembangkan penerapan klasifikasi multi label yang dapat mengelompokkan genre novel berdasarkan gambar sampulnya menggunakan model *Convolutional Neural Network*.
2. Mengumpulkan data set gambar sampul novel dengan teknik *web scraping* yang tepat untuk digunakan dalam melakukan penerapan klasifikasi multi label genre novel.
3. Mengevaluasi hasil pengujian dan kinerja model arsitektur *Convolutional Neural Network* dalam melakukan klasifikasi multi label genre novel berdasarkan gambar sampulnya.

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah sebagai berikut:

1. Sistem dari hasil penelitian ini dapat digunakan untuk mengklasifikasi genre novel berdasarkan gambar sampulnya.
2. Data set sampul novel yang digunakan untuk penerapan klasifikasi multi label yang dikumpulkan oleh penulis dengan teknik *web scraping* dapat berguna dan dipakai untuk penelitian lain.
3. Hasil dari penelitian ini dapat digunakan sebagai bahan kajian dan referensi penelitian dalam melakukan klasifikasi multi label berdasarkan data gambar menggunakan *Convolution Neural Network*.

1.6 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Data *input* yang digunakan dalam penelitian ini berupa data gambar sampul depan novel dengan format JPG dan berukuran 256 x 256.
2. Penelitian ini menggunakan data set sampul depan novel yang dikumpulkan dari *website* FictionDB. Total data yang digunakan ada sebanyak 8.644 data yang terdiri dari 5 kategori genre, yaitu fantasi, misteri, romansa, horor, dan general.
3. Data latih yang digunakan ada 7.500 data, dengan banyak 1.500 data per kategori genre. Data uji yang digunakan ada sebanyak 1.144 data.

1.7 Sistematika Penulisan

Sistematika penulisan tugas akhir ini adalah sebagai berikut:

BAB I. PENDAHULUAN

Pada bab ini diuraikan mengenai latar belakang, perumusan masalah, tujuan dan manfaat penelitian, batasan masalah, dan sistematika penulisan.

BAB II. KAJIAN LITERATUR

Pada bab ini akan membahas dasar-dasar teori yang digunakan dalam penelitian serta beberapa kajian literatur mengenai penelitian lain yang relevan pada penelitian ini.

BAB III. METODOLOGI PENELITIAN

Pada bab ini akan dibahas mengenai tahapan yang akan dilaksanakan pada penelitian ini. Rencana tahapan penelitian dan rancangan hasil pengujian dideskripsikan dengan mengacu pada suatu kerangka kerja.

BAB IV. PENGEMBANGAN PERANGKAT LUNAK

Pada bab ini akan dibahas mengenai proses pengembangan perangkat lunak dari penelitian ini. Semua tahapan mulai dari analisis kebutuhan sistem sampai dengan evaluasi dari sistem yang telah dikembangkan akan diuraikan pada bab ini.

BAB V. HASIL DAN ANALISIS PENELITIAN

Pada bab ini akan dibahas mengenai hasil dari proses pelatihan dan pengujian yang dilakukan dalam penelitian ini. Kemudian dilakukan evaluasi terhadap hasil penelitian yang diperoleh.

BAB VI. KESIMPULAN DAN SARAN

Pada bab ini akan dibahas kesimpulan dari penelitian yang telah dilakukan beserta dengan saran yang diuraikan untuk pengembangan penelitian lebih lanjut di masa yang akan datang.

1.8 Ringkasan

Pada bab 1 ini dapat disimpulkan bahwa masalah yang harus diselesaikan pada penelitian ini adalah bagaimana melakukan klasifikasi multi label genre novel berdasarkan gambar sampul depannya menggunakan metode *Convolutional Neural Network*.

DAFTAR PUSTAKA

- Gibaja, E., & Ventura, S. (2014). Multi-label learning: A review of the state of the art and ongoing research. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 4(6), 411–444. <https://doi.org/10.1002/widm.1139>
- Chiang, H., Ge, Y., & Wu, C. (2015). *Classification of Book Genres By Cover Ana Title*. 2015. https://web.stanford.edu/class/ee368/Project_Autumn_1617/Posters/poster_yang_shen.pdf. [Accessed: 18-Aug-2018].
- O’Shea, K., & Nash, R. (2015). *An Introduction to Convolutional Neural Networks*. 1–11. <http://arxiv.org/abs/1511.08458>
- Courville, I. G. and Y. B. and A. (2016). Deep learning 简介 一 、 什么是 Deep Learning ? . *Nature*, 29(7553), 1–73.
- Novaković, J. D., Veljović, A., Ilić, S. S., Papić, Ž., & Milica, T. (2017). Evaluation of Classification Models in Machine Learning. *Theory and Applications of Mathematics & Computer Science*, 7(1), Pages: 39 – 46. <https://uav.ro/applications/se/journal/index.php/TAMCS/article/view/158>
- Buczowski, P., Sobkowicz, A., & Kozłowski, M. (2018). Deep learning approaches towards book covers classification. *ICPRAM 2018 - Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods*, 2018-January(January), 309–316. <https://doi.org/10.5220/0006556103090316>

- Kulkarni, A., Chong, D., & Batarseh, F. A. (2020). Foundations of data imbalance and solutions for a data democracy. In *Data Democracy: At the Nexus of Artificial Intelligence, Software Development, and Knowledge Engineering*. Elsevier Inc. <https://doi.org/10.1016/B978-0-12-818366-3.00005-8>
- Kundu, C., & Zheng, L. (2020). *Deep multi-modal networks for book genre classification based on its cover*. <http://arxiv.org/abs/2011.07658>
- Shetty, S., Suthar, R., Trivedi, J., Rathod, N., Professor, A., & Gandhi, R. (2020). Convolutional Neural Networks for Genre Classification of Movie Posters. *International Journal for Research in Engineering Application & Management (IJREAM)*, 05(12), 2454–9150. www.imdb.com
- Chollet, F. (2017). *Deep learning with python*. Manning Publications.
- Apis, T., & Tensorflow, B. (n.d.). *3.1 What's TensorFlow?*
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–14.
- Kundalia, K., Patel, Y., & Shah, M. (2019). Multi-label Movie Genre Detection from a Movie Poster Using Knowledge Transfer Learning. *Augmented Human Research*, 0123456789. <https://doi.org/10.1007/s41133-019-0029-y>
- Lydia, A. A., & Francis, F. S. (2020). Multi-Label Classification using Deep Convolutional Neural Network. *International Conference on Innovative*

Trends in Information Technology (ICITIIT). <https://doi.org/10.1088/1742-6596/1584/1/012005>

Sasaki, Y. (2007). The truth of the F-measure. *Teach Tutor Mater*, 1–5.

<http://www.cs.odu.edu/~mukka/cs795sum09dm/Lecturenotes/Day3/F-measure-YS-26Oct07.pdf>

Read, J., & Perez-Cruz, F. (2014). *Deep Learning for Multi-label Classification*. 1–

8. <http://arxiv.org/abs/1502.05988>