

**KLASIFIKASI PENYAKIT KANKER KOLOREKTAL
MENGUNAKAN *CONVOLUTIONAL NEURAL NETWORK*
BERDASARKAN CITRA HISTOLOGI**



TESIS

Diajukan sebagai salah satu syarat untuk memperoleh gelar
Magister Ilmu Komputer (M. Kom)

SITI KHOTIMATUL WILDAH

14002399

Program Studi Ilmu Komputer (S2)

Fakultas Teknologi Informasi

Universitas Nusa Mandiri

2021

SURAT PERNYATAAN ORISINALITAS DAN BEBAS PLAGIARISME

Yang bertanda tangan di bawah ini :

Nama : Siti Khotimatul Wildah
NIM : 14002399
Program Studi : Ilmu Komputer
Fakultas : Teknologi Informasi
Jenjang : Strata Dua (S2)
Konsentrasi : Image Processing

Dengan ini menyatakan bahwa tesis yang telah saya buat dengan judul: “Klasifikasi Penyakit Kanker Kolorektal Menggunakan *Convolutional Neural Network* Berdasarkan Citra Histologi” adalah hasil karya sendiri, dan semua sumber baik yang kutip maupun yang dirujuk telah saya nyatakan dengan benar dan tesis belum pernah diterbitkan atau dipublikasikan dimanapun dan dalam bentuk apapun.

Demikianlah surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila dikemudian hari ternyata saya memberikan keterangan palsu dan atau ada pihak lain yang mengklaim bahwa tesis yang telah saya buat adalah hasil karya milik seseorang atau badan tertentu, saya bersedia diproses baik secara pidana maupun perdata dan kelulusan saya dari Program Studi Ilmu Komputer (S2) Fakultas Teknologi Informasi Universitas Nusa Mandiri dicabut/dibatalkan.

Jakarta, 2 Agustus 2021
Yang menyatakan,



Siti Khotimatul Wildah

HALAMAN PERSETUJUAN DAN PENGESAHAN TESIS

Tesis ini diajukan oleh:

Nama : Siti Khotimatul Wildah
NIM : 14002399
Program Studi : Ilmu Komputer
Jenjang : Strata Dua (S2)
Konsentrasi : *Image Processing*
Judul Tesis : Klasifikasi Penyakit Kanker Kolorektal Menggunakan Convolutional Neural Network Berdasarkan Citra Histologi

Telah dipertahankan pada periode 2021-1 dihadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Magister Ilmu Komputer (M.Kom) pada Program Studi Ilmu Komputer (S2) Fakultas Teknologi Informasi Universitas Nusa Mandiri.

Jakarta, 16 Agustus 2021

PEMBIMBING TESIS

Pembimbing I : Dr. Dewi Yanti Liliana, M.Kom



Pembimbing II : Nita Merlina, M.Kom.



DEWAN PENGUJI

Penguji I : Dr. Rifki Sadikin, M.Kom



Penguji II : Dr. Windu Gata, M.Kom



Penguji III /
Pembimbing I : Dr. Dewi Yanti Liliana, M.Kom





LEMBAR BIMBINGAN TESIS

UNIVERSITAS NUSA MANDIRI

NIM : 14002399
Nama Lengkap : Siti Khotimatul Wildah
Dosen Pembimbing I : Dr. Dewi Yanti Liliana, M.Kom
Judul Tesis : Klasifikasi Penyakit Kanker Kolorektal Menggunakan Convolutional Neural Network Berdasarkan Citra Histologi

No	Tanggal Bimbingan	Pokok Bahasan	Paraf Dosen Pembimbing I
1.	14 April 2021	Pengajuan Judul dan Pengajuan Bab I	
2.	30 April 2021	Revisi Bab I dan Pengajuan Bab II	
3.	7 Mei 2021	Acc Bab I dan Revisi Bab II	
4.	21 Mei 2021	Acc Bab II dan Pengajuan Bab III	
5.	25 Juni 2021	Acc Bab III dan Pengajuan Bab IV	
6.	23 Juli 2021	Acc Bab IV dan Pengajuan Bab V	
7.	30 Juli 2021	Acc Bab V	
8.	2 Agustus 2021	Acc Keseluruhan	

Catatan untuk Dosen Pembimbing.

Bimbingan Tesis

- Dimulai pada tanggal : 14 April 2021
- Diakhiri pada tanggal : 2 Agustus 2021
- Jumlah pertemuan bimbingan : 8 (Delapan) kali bimbingan

Disetujui oleh,
Dosen Pembimbing I

Dr. Dewi Yanti Liliana, M.Kom

KATA PENGANTAR

Puji syukur Alhamdulillah, penulis panjatkan kehadiran Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya sehingga pada akhirnya penulis dapat menyelesaikan tesis yang berjudul “Klasifikasi Penyakit Kanker Kolorektal Menggunakan *Convolutional Neural Network* Berdasarkan Citra Histologi” ini tepat pada waktunya.

Tujuan penulisan laporan tesis ini dibuat sebagai salah satu untuk mendapatkan gelar Magister Ilmu Komputer (M.Kom) pada Program Studi Ilmu Komputer (S2) Universitas Nusa Mandiri.

Laporan Tesis ini diambil berdasarkan hasil penelitian mengenai jenis-jenis kanker kolorektal berdasarkan 5000 citra histologi, penelitian dilakukan mulai dari pengumpulan dataset, proses *reshape* citra, pembagian data *training* dan *testing*, pengklasifikasian menggunakan teknik *deep learning* dengan metode CNN, evaluasi hasil dan implementasi menjadi sebuah sistem. Penulis juga mencari dan menganalisa berbagai macam sumber referensi, baik dalam bentuk jurnal ilmiah, buku-buku literatur, internet, dan lain-lain yang terkait dengan pembahasan pada laporan tesis ini.

Pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada berbagai pihak yang telah memberikan bantuan, dukungan, semangat dan motivasi selama penyusunan tesis, yaitu:

1. Rektor Universitas Nusa Mandiri Ibu Dr. Dwiza Riana, S.Si, M.M, M.Kom
2. Wakil Rektor I Bidang Akademik Universitas Nusa Mandiri
3. Dekan Fakultas Teknologi Informasi Universitas Nusa Mandiri
4. Ketua Program Studi Ilmu Komputer
5. Ibu Dr. Dewi Yanti Liliana, M.Kom dan Ibu Nita Merlina, M.Kom selaku Pembimbing I dan Pembimbing II yang telah meluangkan waktu, tenaga dan pikiran guna memberikan bimbingan bagi penulis selama penyusunan tesis.
6. Seluruh dosen Program Studi Ilmu Komputer (S2) Universitas Nusa Mandiri yang telah memberikan ilmu, pengetahuan dan motivasi selama penulis menempuh studi S2.

7. Yayasan Bina Sarana Informatika yang telah memberikan beasiswa kepada penulis sehingga penulis dapat menempuh pendidikan tinggi hingga S2.
8. Orang tua, keluarga dan sahabat yang telah memberikan dukungan baik secara moral, materiil maupun doa.
9. Teman-teman seperjuangan yang telah memberikan dukungan moral kepada penulis.
10. dan lain-lain.

Serta semua pihak yang tidak dapat penulis sebutkan satu persatu, sehingga atas bantuan semua pihak tersebut penulis dapat menyelesaikan penulisan tesis ini. Penulis menyadari bahwa dalam penulisan tesis ini masih terdapat banyak kekurangan dan kekeliruan, untuk itu penulis mengharapkan kritik dan saran membangun agar penulisan karya ilmiah yang penulis hasilkan di masa yang akan datang menjadi semakin lebih baik.

Akhir kata semoga laporan tesis ini dapat bermanfaat bagi penulis khususnya dan bagi para pembaca yang berminat pada umumnya.

Jakarta, 2 Agustus 2021



Siti Khotimatul Wildah

Penulis

SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Yang bertanda tangan di bawah ini, saya :

Nama : Siti Khotimatul Wildah

NIM : 14002399

Program Studi : Ilmu Komputer

Fakultas : Teknologi Informasi

Jenjang : Strata Dua (S2)

Konsentrasi : Image Processing

Jenis Karya : Tesis

Demi pengembangan ilmu pengetahuan, dengan ini menyetujui untuk memberikan ijin kepada pihak Program Studi Ilmu Komputer (S2) Universitas Nusa Mandiri **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalti-Free Right*)** atas karya ilmiah kami yang berjudul : "Klasifikasi Penyakit Kanker Kolorektal Menggunakan *Convolutional Neural Network* Berdasarkan Citra Histologi" beserta perangkat yang diperlukan (apabila ada).

Dengan **Hak Bebas Royalti Non-Eksklusif** ini pihak Universitas Nusa Mandiri berhak menyimpan, mengalih-media atau *bentuk*-kan, mengelolanya dalam pangkalan data (*database*), mendistribusikannya dan menampilkan atau mempublikasikannya di *internet* atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari kami selama tetap mencantumkan nama kami sebagai penulis/pencipta karya ilmiah tersebut.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Universitas Nusa Mandiri, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 2 Agustus 2021

Yang Menyatakan



Siti Khotimatul Wildah

ABSTRAK

Nama : Siti Khotimatul Wildah
NIM : 14002399
Program Studi : Ilmu Komputer
Fakultas : Teknologi Informasi
Jenjang : Strata Dua (S2)
Konsentrasi : Image Processing
Judul Tesis : “Klasifikasi Penyakit Kanker Kolorektal Menggunakan *Convolutional Neural Network* Berdasarkan Citra Histologi”

Kanker Kolorektal merupakan perkembangan dari keganasan kanker yang terjadi pada jaringan usus besar atau rektum dan berada pada urutan ketiga jenis kanker ganas di dunia dengan rasio sebesar 5% yaitu 1 dari 20 orang. Kanker kolorektal dapat terjadi karena adanya riwayat polip adenoma, riwayat penyakit kronis inflamatori yang terjadi pada usus, kurangnya aktifitas fisik, kelebihan berat badan, konsumsi daging atau lemak yang tinggi, perokok aktif, dan mengonsumsi alkohol. Pada tahap awal kanker kolorektal sering kali tidak menimbulkan gejala, dan baru disadari setelah tumor tumbuh cukup besar, oleh sebab itu pentingnya tindakan pemeriksaan awal agar dapat mencegah kanker tersebut semakin memburuk dan terhindar dari berbagai komplikasi. Penelitian ini bertujuan untuk mengusulkan metode yang dapat mengklasifikasikan penyakit kanker kolorektal dengan melakukan pengujian terhadap data hasil proses biopsi sebanyak 5.000 citra histologi. Metode yang diusulkan pada penelitian ini menggunakan teknik *deep learning* dengan metode CNN dan nilai akurasi yang diperoleh sebesar 89,60%.

Kata Kunci : Kanker Kolorektal, Citra Histologi, *Deep Learning*, CNN

ABSTRACT

Name : Siti Khotimatul Wildah
NIM : 14002399
Study of Program : Ilmu Komputer
Faculty : Information Technology
Levels : Strata Dua (S2)
Concentration : Image Processing
Title : “*Classification of Colorectal Cancer Using Convolutional Neural Network Based on Histological Image*”

Colorectal cancer is a development of malignancy of cancer that occurs in the colon or rectum tissue and is ranked third malignant cancer in the world with a ratio of 5% which is 1 in 20 people. Colorectal cancer can occur due to a history of adenoma polyps, a history of chronic inflammatory diseases that occur in the intestines, lack of physical activity, overweight, high consumption of meat or fat, active smokers, and alcohol consumption. In the early stages of colorectal cancer often does not cause symptoms, and only realized after the tumor grows large enough, therefore the importance of early screening measures in order to prevent the cancer is getting worse and avoid various complications. This study aims to propose a method that can classify colorectal cancer by testing the data of the biopsy process as many as 5.000 histological images. The method proposed in this study uses deep learning techniques with CNN method and accuracy value obtained by 89,60%.

Keywords: Colorectal Cancer, Histology Image, Deep Learning, CNN

DAFTAR ISI

	Halaman
HALAMAN JUDUL.....	i
SURAT PERNYATAAN ORISINALITAS DAN BEBAS PLAGIARISM.....	ii
PERSETUJUAN DAN PENGESAHAN.....	iii
LEMBAR BIMBINGAN TESIS	iv
KATA PENGANTAR	v
SURAT PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS	vii
ABSTRAK.....	viii
ABSTRACT.....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang Penulisan	1
1.2. Identifikasi Masalah	3
1.3. Tujuan Penelitian.....	3
1.4. Ruang Lingkup Penelitian	4
1.5. Hipotesis	4
1.6. Sistematika Penulisan.....	4
BAB II LANDASAN/ KERANGKA PEMIKIRAN	6
2.1. Tinjauan Pustaka	6
2.1.1. Kanker Kolorektal	6
2.1.2. Klasifikasi.....	7
2.1.3. Citra Histologi	7
2.1.4. <i>Deep Learning</i>	7
2.1.5. CNN (<i>Convolutional Neural Network</i>)	7
2.1.6. <i>Confusion Matrix</i>	10
2.1.7. <i>Optimizer</i>	10
2.1.8. <i>Python</i>	11
2.1.9. <i>Flask Python</i>	12
2.1.10. <i>Google Colab</i>	12
2.2. Tinjauan Studi	12
2.3. Tinjauan Organisasi/ Obyek Penelitian	20
BAB III METODOLOGI PENELITIAN	21
3.1. Tahapan Penelitian	21
3.1.1. Pemanggilan Citra	22
3.1.2. Pra Pemrosesan.....	24
3.1.3. Pembagian Dataset	24
3.1.4. Pembangunan Model CNN	24

	3.1.5. Penentuan <i>Hyperparameter</i>	27
	3.1.6. Pelatihan Model.....	27
	3.1.7. Evaluasi Model.....	27
	3.1.8. Implementasi Sistem	28
	3.2. Instrumen Penelitian.....	28
	3.3. Metode Pengumpulan Data	29
BAB IV	HASIL PENELITIAN DAN PEMBAHASAN	30
	4.1. Skenario Uji Coba	30
	4.1.1. Hasil Uji Coba Metode Usulan	30
	4.1.2. Hasil Uji Coba Jumlah Layer Konvolusi	32
	4.1.3. Hasil Uji Coba <i>Epoch</i>	35
	4.1.4. Hasil Uji Coba Persentase <i>Dataset</i>	36
	4.1.5. Hasil Uji Coba Jenis Optimasi	37
	4.1.6. Hasil Uji Coba Metode <i>Transfer Learning</i>	38
	4.2. Evaluasi Hasil Klasifikasi.....	38
	4.3. Implementasi Sistem	44
BAB V	PENUTUP	47
	5.1. Kesimpulan.....	47
	5.2. Saran	48
	DAFTAR REFERENSI	49
	DAFTAR RIWAYAT HIDUP	52
	LAMPIRAN LAMPIRAN	53

DAFTAR TABEL

	Halaman
Tabel 2. 1 <i>Confusion Matrix</i>	10
Tabel 2. 2 Rangkuman Penelitian Terkait.....	16
Tabel 3. 1 Sample Citra Dataset Kanker Kolorektal.....	21
Tabel 3. 2 Arsitektur Usulan CNN.....	24
Tabel 3. 2 Arsitektur Usulan CNN.....	24
Tabel 4. 1 Arsitektur CNN 1 Layer Konvolusi	31
Tabel 4. 2 Arsitektur CNN 2 Layer Konvolusi	31
Tabel 4. 3 Arsitektur CNN 3 Layer Konvolusi	32
Tabel 4. 4 Arsitektur CNN 4 Layer Konvolusi	32
Tabel 4. 5 Arsitektur CNN 5 Layer Konvolusi	32
Tabel 4. 6 Nilai Akurasi Berdasarkan Jumlah Layer Konvolusi	34
Tabel 4. 7 Nilai Akurasi Berdasarkan Jumlah <i>Epoch</i>	35
Tabel 4. 8 Nilai Akurasi Berdasarkan Persentase Data Uji dan Data Latih.....	35
Tabel 4. 9 Nilai Akurasi Berdasarkan Jenis Optimasi	36
Tabel 4. 10 Nilai Akurasi <i>Transfer Learning</i>	37
Tabel 4. 11 <i>Confusion Matrix</i>	40
Tabel 4. 12 Laporan Klasifikasi	43

DAFTAR GAMBAR

	Halaman
Gambar 3. 1 Tahapan Penelitian	20
Gambar 3. 2 Model CNN	25
Gambar 3. 3 Rancangan Antarmuka Aplikasi.....	27
Gambar 4. 1 Usulan Model Arsitektur CNN	30
Gambar 4. 2 Hasil Prediksi Citra Jaringan Kanker Kolorektal.....	38
Gambar 4. 3 Grafik Model Usulan.....	39
Gambar 4. 4 Hasil Grafik Akurasi dan <i>Loss</i>	40
Gambar 4. 5 Tampilan Awal Program	44
Gambar 4. 6 Pemilihan Citra.....	44
Gambar 4. 7 Proses Input Citra.....	45
Gambar 4. 8 Hasil Prediksi Kelas Jaringan Adipose	45
Gambar 4. 9 Hasil Prediksi Kelas Jaringan Mucosa.....	46

DAFTAR LAMPIRAN

	Halaman
Lampiran 1. Dataset Citra Histologi Kanker Kolorektal	53
Lampiran 2. Tampilan Antarmuka Program	57
Lampiran 3. Coding Pengolahan dan Pemodelan CNN.....	61

BAB I

PENDAHULUAN

Bab ini berisi penjelasan dari beberapa hal yang mendasari dilakukannya penelitian yang meliputi, latar belakang penelitian, identifikasi permasalahan yang diangkat, tujuan dilakukannya penelitian dimana tujuan ini menjawab dari permasalahan yang ada, ruang lingkup atau cakupan penelitian yang dilakukan agar penelitian memiliki konsep yang jelas dan terperinci, hipotesis atau dugaan sementara terhadap penelitian yang dilakukan serta sistematika penulisan.

1.1. Latar Belakang Masalah

Kanker Kolorektal merupakan perkembangan dari keganasan kanker yang terjadi pada jaringan usus besar atau rektum (bagian kecil dari usus besar sebelum anus) [1]. Kanker kolorektal berada pada urutan ketiga jenis kanker ganas di dunia, dan diperkirakan pada tahun 2030 akan meningkat sebesar 60%, dari 2,2 juta kasus baru dengan prevalansi 1,1 juta kematian akibat kanker tersebut [2].

Rasio dari risiko keseluruhan kanker kolorektal sebesar 5% yaitu 1 dari 20 orang serta seringkali laki-laki lebih berisiko mengidap penyakit kanker tersebut dibandingkan dengan wanita, hal ini dapat terjadi karena adanya riwayat polip adenoma individual atau keluarga, riwayat penyakit kronis inflamatori yang terjadi pada usus, kurangnya aktifitas fisik, kelebihan berat badan, konsumsi daging atau lemak yang tinggi, perokok aktif, dan mengonsumsi alkohol (Kementerian Kesehatan Republik Indonesia, 2014). Secara umum faktor penyebab terjadinya kanker kolorektal merupakan riwayat penyakit pada usus dan pola hidup yang tidak sehat. Pada tahap awal kanker kolorektal sering kali tidak menimbulkan gejala, dan baru disadari setelah tumor tumbuh cukup besar, oleh sebab itu pentingnya tindakan pemeriksaan awal agar dapat mencegah kanker tersebut semakin memburuk dan terhindar dari berbagai komplikasi.

Pentingnya deteksi dini diperlukan pada kasus ini, agar dapat menentukan perkembangan dari kanker kolorektal yang berpengaruh terhadap tindak lanjut yang akan dilakukan dalam pengobatannya sehingga dapat menghentikan laju pertumbuhan kanker maupun pencegahan tahap awal penyakit tersebut. Diagnosis

Kanker Kolorektal dapat dilakukan melalui foto *rontgen*, MRI, kolonoskopi, biopsi, *ct-scan*, maupun *ultrasound*. Teknik biopsi merupakan salah satu prosedur yang baik untuk dapat menentukan dan mendiagnosis kelainan yang ada pada kolon atau rektum melalui citra histologi [3]. Hasil diagnosis berdasarkan citra histologi tersebut kemudian dianalisis oleh dokter maupun tenaga medis sehingga dapat ditentukan jenis kanker kolorektar yang dialami, akan tetapi pengklasifikasian jenis kanker kolorektal secara manual memerlukan waktu yang relatif lama, oleh sebab itu diperlukannya sebuah teknologi yang dapat memudahkan dokter maupun tenaga medis dalam pengklasifikasian jenis kanker kolorektal tersebut.

Pada saat ini perkembangan teknologi informasi memungkinkan adanya pengambilan informasi pada sebuah citra berdasarkan ciri citra tersebut secara cepat [4]. Pengambilan informasi atau analisa informasi pada citra dapat dilakukan dengan memanfaatkan teknologi *computer vision*, sehingga diagnosis dan penanganan penyakit kanker kolorektal dapat dilakukan lebih cepat, akan tetapi dalam penerapan teknologi *computer vision* diperlukan sebuah metode atau algoritma yang baik sehingga dapat memprediksi penyakit kanker kolorektal dengan akurat. Penentuan model maupun algoritma dapat berpengaruh pada hasil prediksi yang diperoleh, maka pada penelitian ini akan dilakukan penerapan teknik *deep learning* dengan menggunakan metode CNN.

Penelitian ini mengusulkan metode dengan menggunakan *Convolutional Neural Network* (CNN) yang termasuk kedalam salah satu implementasi dari teknologi *deep learning* dan merupakan metode populer untuk pemrosesan citra digital karena berada pada urutan pertama kategori *Image Classification* mulai tahun 2012 sampai dengan tahun 2015 [5]. Metode CNN juga terbukti memiliki performa akurasi yang baik dalam memprediksi data penyakit seperti pada penelitian [6] mengenai penerapan metode CNN untuk melakukan diagnosa tumor otak berdasarkan citra MRI. Pada penelitian tersebut dinyatakan bahwa metode CNN memiliki performa akurasi rata-rata diatas 90%. Kemudian penelitian yang dilakukan oleh [7] mengenai penerapan *deep learning* menggunakan metode CNN untuk melakukan klasifikasi penyakit *stroke* pada citra *CT-Scan* dengan perolehan akurasi sebesar 96,78%.

Penelitian ini dilakukan dengan menguji dataset yang didapatkan dari situs *repository kaggle* dan berasal dari arsip patologi Institut Patologi, Pusat Medis Universitas Mannheim, Universitas Heidelberg, Mannheim, Jerman dengan jumlah data sebanyak 5.000 citra histologi jenis jaringan Kanker Kolorektal dengan menggunakan algoritma CNN.

1.2. Identifikasi Masalah

Berdasarkan latar belakang diatas, terdapat permasalahan yang dapat diidentifikasi pada penelitian ini diantaranya:

1. Bagaimana menerapkan model CNN dengan arsitektur yang telah dimodifikasi pada pengklasifikasian penyakit kanker kolorektal berdasarkan citra histologi?
2. Bagaimana menerapkan *hyperparameter* yang baik agar model CNN yang diusulkan dapat bekerja dengan baik?
3. Bagaimana mengukur atau mengevaluasi seberapa akurat metode CNN yang diusulkan?
4. Bagaimana mengimplementasikan metode CNN yang diusulkan pada pengklasifikasian penyakit kanker kolorektal?

1.3. Tujuan Penelitian

Penelitian ini mempunyai tujuan yang diharapkan dapat terlaksana dengan baik. Tujuan dari penelitian ini, antara lain:

1. Melakukan pemodelan dengan metode CNN menggunakan arsitektur yang telah dilakukan modifikasi baik layer konvolusi, jumlah filter, ukuran kernel maupun *flatten layer* untuk mengklasifikasikan penyakit kanker kolorektal.
2. Melakukan penerapan *hyperparameter* yang baik dengan tujuan untuk mendapatkan model CNN yang diusulkan dapat bekerja dengan baik.
3. Mengukur dan melakukan evaluasi tingkat keakuratan model CNN yang diusulkan.
4. Melakukan implementasi metode CNN pada pengklasifikasian penyakit kanker kolorektal.

1.4. Ruang Lingkup Penelitian

Agar masalah yang dibahas fokus terhadap materi yang akan diteliti maka penelitian ini dibatasi dengan ruang lingkup pengujian model algoritma CNN (*Convolutional Neural Network*) untuk mengklasifikasikan penyakit kanker kolorektal berdasarkan citra histologi. Dataset yang digunakan diambil dari situs *repository kaggle* dan berasal dari arsip patologi Institut Patologi, Pusat Medis Universitas Mannheim, Universitas Heidelberg, Mannheim, Jerman sebanyak 5000 citra histologi dengan 8 jenis kelas yang terdiri dari *tumor, stroma, lympho, debris, mucosa, adipose, non-tumor (empty)* sampai dengan kanker kolorektal yang kompleks. Pengolahan citra menggunakan bahasa pemrograman *Python* dengan *tools* yang digunakan *Google Collaboratory*. Evaluasi model dilakukan dengan mengukur nilai akurasi, *recall, precision, confusion matrix* dan grafik akurasi. Implementasi model dibuat dalam bentuk program berbasis *website* dengan *framework Python Flask*.

1.5. Hipotesis

Hipotesis atau dugaan awal pada penelitian klasifikasi penyakit kanker kolorektal berdasarkan citra histologi diantaranya:

H₁: Diduga pemodelan CNN dengan arsitektur yang diusulkan dapat mendiagnosa penyakit kanker kolorektal dengan baik.

H₂: Diduga penerapan beberapa *hyperparameter* dapat mempengaruhi kinerja model

H₃: Diduga pemodelan CNN yang diusulkan dapat diimplementasikan dengan baik pada pengklasifikasian kanker kolorektal.

1.6. Sistematika Penulisan

Sistematika penulisan bertujuan untuk memberikan konsep mengenai permasalahan yang diangkat.

BAB I PENDAHULUAN

Bab ini berisi penjelasan tentang latar belakang penulisan, identifikasi masalah, tujuan penelitian, ruang lingkup penelitian dan

hipotesis, serta sistematika penulisan dari penelitian yang akan disusun.

BAB II LANDASAN TEORI

Bab ini menjelaskan tinjauan pustaka dari teori-teori yang berhubungan dengan penelitian yang dilakukan, tinjauan studi dari penelitian-penelitian sebelumnya yang membahas mengenai permasalahan yang serupa, serta tinjauan obyek yang digunakan dalam penelitian.

BAB III METODE PENELITIAN

Bab ini membahas mengenai metode pengumpulan data, perencanaan pembangunan model dari metode penelitian yang dipilih serta konsep dari penelitian yang bertujuan untuk menyelesaikan pokok permasalahan yang diteliti.

BAB IV HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil penelitian, berupa hasil pengolahan citra, hasil pengujian, perhitungan manual, implementasi sistem dan pembahasan.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari hasil penelitian dan saran-saran yang dibutuhkan untuk penelitian lebih lanjut dimasa yang akan datang.

BAB II

LANDASAN TEORI

Bab ini berisi pembahasan mengenai referensi yang mendukung dan mendasari penelitian yang dilakukan dengan tujuan sebagai pendukung dalam menyelesaikan permasalahan yang telah diuraikan pada latar belakang, dimana pada bab ini berisi mengenai tinjauan pustaka, penelitian terkait yang telah dilakukan sebelumnya, serta penjelasan mengenai objek penelitian yang akan digunakan.

2.1. Tinjauan Pustaka

Tinjauan pustaka berisi mengenai teori-teori serta metode yang digunakan dalam penelitian. Referensi terkait penelitian diambil dari buku, jurnal nasional dan internasional yang berhubungan dengan judul penelitian.

2.1.1. Kanker Kolorektal

Kanker Kolorektal merupakan perkembangan dari keganasan kanker yang terjadi pada jaringan usus besar atau rektum (bagian kecil dari usus besar sebelum anus) [1]. Kanker kolorektal berada pada urutan ketiga jenis kanker ganas di dunia, dan diperkirakan pada tahun 2030 akan meningkat sebesar 60%, dari 2,2 juta kasus baru dengan prevalansi 1,1 juta kematian akibat kanker tersebut [2].

Beberapa faktor risiko yang dapat meningkatkan timbulnya penyakit kanker kolorektal diantaranya faktor genetik seperti riwayat keluarga penderita kanker kolorektal, usia, diet tidak sehat [8], kurangnya aktifitas fisik, kelebihan berat badan, konsumsi daging atau lemak yang tinggi, perokok aktif, dan mengonsumsi alkohol (Kementerian Kesehatan Republik Indonesia, 2014).

Gejala umum yang terjadi pada penderita kanker kolorektal yakni buang air besar (BAB) yang berlebih dan juga berdarah, nyeri pada perut, kondisi badan yang lemah dan lesu, serta penurunan berat badan [8].

2.1.2. Klasifikasi

Klasifikasi citra merupakan proses pengelompokan semua piksel atau representasi sebuah titik terkecil dalam suatu citra sehingga dapat didefinisikan atau ditafsirkan sebagai atribut tertentu secara spesifik [9].

2.1.3. Citra Histologi

Citra Histologi berfungsi sebagai acuan untuk diagnosis dari suatu penyakit karena dari citra histologi ini dapat dilihat perubahan fungsi fisiologi dan perubahan bentuk dari suatu organ [10]. Citra histologi didapat melalui pemeriksaan jaringan utuh melalui teknik biopsi. Teknik biopsi merupakan salah satu prosedur yang baik untuk dapat menentukan dan mendiagnosis kelainan yang ada pada kolon atau rektum [3].

2.1.4. Deep Learning

Deep learning merupakan suatu metode pembelajaran dengan menggunakan jaringan syaraf tiruan *multi layer*, yang mirip seperti otak manusia, dimana *neuron-neuron* saling terhubung membentuk jaringan *neuron* yang sangat kompleks [11]. *Deep learning* merupakan pengembangan dari kecerdasan buatan untuk pembelajaran mendalam, memiliki konsep yang sama seperti *neural network* namun tidak memerlukan deskripsi fitur. Dalam algoritma *deep learning* terdapat proses input, kemudian proses ekstraksi fitur yang sudah dilakukan secara bersamaan dengan proses klasifikasi [12].

2.1.5. CNN (Convolutional Neural Network)

CNN merupakan operasi konvolusi yang menggabungkan beberapa lapisan pemrosesan, menggunakan beberapa elemen operasi paralel, dan terinspirasi dari sistem saraf biologis. Pada CNN, setiap *neuron* disajikan dalam bentuk dua dimensi, sehingga algoritma ini cocok untuk memproses data berbentuk citra [13]. CNN termasuk kedalam metode *Deep Learning* dikarenakan pada pengaplikasiannya memiliki kedalaman jaringan yang tinggi dan CNN banyak diaplikasikan untuk menganalisis masukan berupa citra visual, analisa dan deteksi objek pada sebuah citra [11].

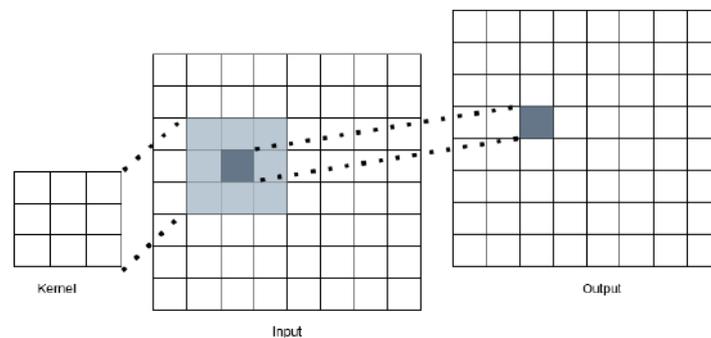
CNN bekerja menggunakan kernel. Kernel tersebut akan mengekstrak fitur dari input menggunakan operasi konvolusi [14]. Arsitektur CNN terdiri 5 bagian utama yaitu *input layer*, *convolution layer*, *pooling layer*, *fully-connected layer* dan *output layer* [5].

1.1. *Input Layer*

Input layer merupakan input data-data citra yang dikonversi menjadi matriks dengan 3 dimensi dimana nilai masing-masing dimensi yaitu warna merah, biru dan hijau [5]. Penggabungan dari seluruh matriks peta fitur yang diperoleh dari proses penggabungan lapisan, dan kemudian semua piksel sepanjang jumlah piksel dari matriks yang diperoleh pada lapisan penggabungan direntangkan menjadi vektor. Setelah penggabungan menjadi *vector*, maka gunakan semua nilai di lapisan masukan untuk dikalkulasikan pada lapisan tersembunyi (*hidden layer*).

1.2. *Convolution Layer*

Merupakan bagian utama dari CNN, operasi yang dilakukan sama dengan operasi konvolusi yang biasa dilakukan pada pengolahan citra (*image processing*), yang terdiri dari kernel dan sub citra dengan ukuran kernel yang umumnya digunakan adalah 3x3. Dimana pada setiap sub citra dan kernel yang berukuran sama dilakukan operasi konvolusi [11] seperti pada gambar 2.1.



Gambar 2.1 Convolutional Layer

Pada lapisan ini, operasi konvolusi akan dilakukan antara matriks citra masukan dan matriks *filter*. *Filter* tersebut akan dipindahkan ke seluruh permukaan citra, sehingga menghasilkan keluaran matriks peta fitur, peta fitur yang dihasilkan dapat diperoleh dari persamaan berikut.

$$n_{out} = \left(\frac{n_{in} - k + 2p}{s} \right) + 1 \quad (1)$$

dimana:

n_{out} : Ukuran *feature map*

n_{in} : Ukuran matriks masukan

k : Ukuran matriks *filter*

p : Ukuran *padding*

s : *Stride*

Adapun persamaan operasi konvolusi adalah sebagai berikut.

$$FM[i]_{j,k} = \left(\sum_m \sum_n N_{[j-m,k-n]} F_{[m,n]} \right) + Bf \quad (2)$$

dimana :

$FM[i]$: Matriks *Feature Map* ke-i

N : Matriks citra masukan

F : Matriks *filter* konvolusi

bF : Nilai bias pada *filter*

j,k : Posisi piksel pada matriks citra masukan

m,n : Posisi piksel pada matriks filter konvolusi

Setelah pemrosesan konvolusi selesai, *Rectified Linear Unit* (ReLU) kemudian digunakan untuk mengaktifkan fungsi aktivasi. Setiap piksel dalam peta fitur akan dimasukkan ke fungsi ReLU, dan piksel dengan nilai kurang dari 0 akan diubah menjadi 0, dengan persamaan sebagai berikut.

$$f(x) = \max(0, x) \quad (3)$$

1.3. Pooling Layer

Pooling layer tersusun atas sebuah *filter* yang berukuran dan memiliki *stride* tertentu, dimana setiap perpindahan yang dilakukan berdasarkan oleh jumlah *stride* yang berpindah pada seluruh area *feature map* atau *activation map*. Adapun implementasi dari *pooling layer* ini biasanya menggunakan jenis *max pooling* dan *average pooling*.

1.4. Flatten Layer

Flatten layer merupakan layer yang memiliki proses bahwa setiap bari yang ada pada *feature map* akan disusun secara vertikal dan tujuan dari proses pada *flatten layer* ini adalah mengubah data dua dimensi menjadi satu dimensi [15].

1.5. Fully Connected Layer

Merupakan proses tahap klasifikasi *multilayer perceptron* (MLP). Pada lapisan *fully connected layer* tersusun atas setiap *neurons* yang memiliki hubungan penuh terhadap seluruh aktivasi pada lapisan sebelumnya.

1.6. Dropout Layer

Dropout merupakan salah satu teknik yang digunakan untuk mencegah terjadinya *overfitting* serta digunakan untuk mempercepat proses *learning*. *Dropout* bekerja dengan menghilangkan sementara suatu neuron *hidden layer* maupun *visible layer* yang berada pada jaringan [11].

2.1.6. Confusion Matrix

Confusion matrix merupakan metode yang digunakan untuk mengevaluasi kinerja suatu algoritma klasifikasi dengan menggunakan table matrik [16]. Dalam *confusion matrix* *True Positive* (TP) dan *True Negative* (TN) merupakan jumlah kelas positif dan negatif yang diklasifikasikan dengan tepat, *False Positive* (FP) dan *False Negative* (FN) merupakan jumlah kelas positif dan negatif yang tidak diklasifikasikan dengan tepat [17].

Tabel 2.1 Confusion Matrix

<i>Class</i>	<i>Predictive Positive</i>	<i>Predictive Negative</i>
<i>Actual Positive</i>	<i>True Positive</i> (TP)	<i>False Positive</i> (FP)
<i>Actual Negative</i>	<i>False Negative</i> (FN)	<i>True Negative</i> (TN)

Sumber: Siringoringo, 2018

2.1.7. Optimizer

Optimizer adalah metode atau algoritma yang digunakan untuk memperbarui bobot, dengan tujuan untuk mengurangi nilai kerugian atau *loss*, sehingga menghasilkan skor akurasi yang baik. *Optimizer* digunakan untuk meminimalkan kesalahan, memaksimalkan akurasi serta menemukan bobot terbaik,

dimana pada saat terjadinya proses pelatihan, parameter model dapat diubah dengan tujuan untuk meminimalkan nilai *loss* agar prediksi yang dihasilkan akurat. [18].

Berikut beberapa optimizer yang sering digunakan diantaranya:

1. *Adaptive Moment Estimation* (Adam) merupakan metode yang menghitung *adaptive learning rate* untuk setiap parameter. Nilai parameter yang direkomendasikan adalah $\beta_1 = 0.9$, $\beta_2 = 0.999$, dan $\epsilon = 10^{-8}$. Dengan $\beta_1 = \beta_2 =$ tingkat penurunan eksponensial dan $\epsilon =$ nilai epsilon untuk update parameter [19].
2. *RMSprop* merupakan metode yang mempertahankan *moving average* dari gradien kuadrat untuk setiap bobot. Nilai parameter yang direkomendasikan adalah $\eta = 0.001$ dengan $\eta =$ *learning rate* dengan *bias-correction* [18].
3. *Stochastic Gradient Descent* (SGD) merupakan algoritma yang digunakan untuk memperbaharui bobot parameter dan deviasi, Algoritma ini pada dasarnya hanya mengurangi bobot awal atau bagian dari nilai gradien yang sudah diperoleh [20].
4. *Adagrad* merupakan varian dari algoritma *Stochastic Gradient Descent* (SGD) yang mengubah tingkat pembelajarannya menjadi parameter yang berbeda. Algoritma ini diterbitkan pada tahun 2011, oleh John Duchi, dimana algoritma ini mengadaptasi tingkat pembelajaran yang lebih besar untuk data yang lebih sedikit dan lebih kecil akan tetapi untuk parameter yang berkala [21].
5. *Adadelta* merupakan metode lanjutan dari *Adagrad* yang berusaha mengurangi tingkat pembelajaran yang agresif dan menurun secara monoton. Metode ini bekerja dengan membatasi jendela gradien masa lalu yang terakumulasi untuk beberapa ukuran tetap w [22].

2.1.8. Python

Python merupakan bahasa pemrograman interpretatif berorientasi objek yang memiliki kode-kode pemrograman yang jelas, lengkap dan mudah untuk dipahami serta berorientasi objek [23]. Bahasa pemrograman *Python* menyediakan banyak struktur data dan fungsi yang berguna untuk data science [24].

Python dikembangkan oleh Guido van Rossum seorang *programmer* kelahiran Belanda di CWI di Amsterdam pada tahun 1990, dimana bahasa pemrograman *python* merupakan kelanjutan dari bahasa pemrograman ABC. Nama *Python* diambil dikarenakan Guido menyukai acara TV *Monty Python's Flying Circus* [23]. Penelitian ini menggunakan *Python* versi 3.

2.1.9. Flask Python

Flask merupakan kerangka kerja web yang ditulis dengan bahasa pemrograman *Python* dan diklasifikasikan sebagai kerangka mikro. *Flask* bertindak sebagai kerangka aplikasi dan tampilan dengan nuansa Web. Dengan menggunakan bahasa *Flask* dan *Python*, pengembang dapat membuat Web terstruktur dan mengelola perilaku Web dengan lebih mudah [25]. *Flask* termasuk dalam jenis *micro-framework* karena tidak memerlukan *tools* atau *library* tertentu dalam penggunaannya.

2.1.10. Google Colab

Google Colab merupakan *tools* yang dikembangkan oleh *Google Internal Research* dan merupakan salah satu produk *Google* yang berbasis *cloud* dan bersifat *open source*. *Google Colab* merupakan editor dengan bahasa pemrograman *Python* dengan format *notebook*. *Google Colab* memfasilitasi layanan GPU secara gratis dengan pelayanan selama 12 jam dalam suatu waktu, bersifat *cloud* dan menyimpan berkas kedalam *Google Drive* [26].

2.2. Tinjauan Studi

Penelitian mengenai penyakit kanker kolorektal telah dilakukan sebelumnya menggunakan teknik pengolahan citra. Berikut beberapa penelitian yang dijadikan sebagai acuan pada penelitian ini:

1. Penelitian yang dilakukan oleh Lavita Nuraviana Rizalputri, Timothy Pranata, Nancy Silvia Tanjung, Hasna Marhamah Auliya, Suksmandhira Harimurti, dan Isa Anshori pada tahun 2019 mengenai perbandingan akurasi multiklasifikasi histologi kolorektal menggunakan berbagai model *machine learning*, dimana *dataset* yang digunakan berasal dari University Medical

Center Mannheim, Heidelberg University, Mannheim, Jerman. Metode yang digunakan pada penelitian ini diantaranya *Convolutional Neural Network* (CNN), *K-Nearest Neighbour* (KNN), *Logistic Regression*, dan *Random Forest*. Pada penelitian ini metode CNN memberikan hasil terbaik dengan nilai akurasi yang dihasilkan sebesar 82,2% sedangkan KNN sebesar 62,56%, RF sebesar 68,72% dan LR sebesar 52,6% [27].

2. Penelitian yang dilakukan oleh Kun Zhang , JunHong Fu, Liang Hua , Peijian Zhang, Yeqin Shao, Sheng Xu, Huiyu Zhou, Li Chen, dan Jing Wang pada tahun 2016 mengenai analisis citra patologi kanker kolorektal dengan menggunakan *dataset* MICCAI GlaS 2015 (Nantong University). Metode yang diusulkan pada penelitian ini menggunakan CNN, dengan hasil yang diperoleh sebesar 86,91% [28].
3. Penelitian yang dilakukan oleh Namita Sengar, Neeraj Mishra, Malay Kishore Dutta, Jiri Prinosil, dan Radim Burget pada tahun 2016 mengenai penilaian kanker kolorektal menggunakan citra histologi. Metode yang diusulkan pada penelitian ini menggunakan SVM, dengan menggunakan *dataset* *Warwick-QU dataset*. Hasil yang diperoleh pada penelitian ini yaitu akurasi sebesar 80%, sensitivitas 87%, dan spesifisitas 83% [29].
4. Penelitian yang dilakukan oleh Devvi Sarwindaa, Radifa Hilya Paradisaa, Alhadi Bustamama dan Pinkie Anggiab pada tahun 2021 mengenai penerapan teknik *Deep Learning* dalam klasifikasi citra histologi menggunakan *varian residual network* (ResNet). *Dataset* yang digunakan pada penelitian ini adalah *Warwick-QU dataset*. Adapun hasil yang diperoleh yaitu Akurasi sebesar 80%, sensitivitas 87%, dan spesifisitas 83% [30].
5. Penelitian yang dilakukan oleh Jakob Nikolas Kather, Cleo-AronWeis, Francesco Bianconi, Susanne M. Melchers, Lothar R. Schad, TimoGaiser, Alexander Marx dan FrankGerritZöllner pada tahun 2016 mengenai analisis tekstur multi-kelas dalam histologi kanker kolorektal. Penelitian ini menyajikan kumpulan data dari 5.000 citra histologis kanker kolorektal manusia termasuk delapan jenis jaringan yang berbeda, dimana *dataset* yang digunakan diperoleh dari Pusat Medis Universitas Mannheim, Universitas Heidelberg, Jerman. Metode yang diusulkan pada penelitian ini

- menggunakan metode SVM dengan akurasi yang diperoleh sebesar 87,4% [31].
6. Penelitian yang dilakukan oleh Silvia Cascianelli, Raquel Bello-Cerezo, Francesco Bianconi, Mario L. Fravolini, Mehdi Belal, Barbara Palumbo, dan Jakob N pada tahun 2018 mengenai strategi pengurangan dimensi untuk klasifikasi citra histopatologi. Penelitian ini menggunakan metode CNN dengan *dataset* yang digunakan adalah dataset yang diperoleh dari Pusat Medis Universitas Mannheim, Universitas Heidelberg, Jerman. Hasil yang diperoleh pada penelitian ini sebesar 75% [32].
 7. Penelitian yang dilakukan oleh Mary Shapcott, Katherine J. Hewitt dan Nasir Rajpoot pada tahun 2019 mengenai penerapan teknik *deep learning* dengan pengambilan sampel dalam histologi kanker usus besar. Penelitian ini menggunakan *dataset The Cancer Genome Atlas (TCGA)* dengan metode yang diusulkan yaitu CNN dan hasil akurasi yang diperoleh sebesar 84% [33].
 8. Penelitian yang dilakukan oleh Mindaugas Morkunas, Povilas Treigys, Jolita Bernatavičiene, Arvydas Laurinavičius, dan Gražina Korvel pada tahun 2018 mengenai klasifikasi jaringan tumor kanker kolorektal berbasis *machine learning*. Penelitian ini menggunakan *dataset National Center of Pathology, Lithuania* dengan metode yang diusulkan yaitu CNN dan hasil akurasi yang diperoleh pada penelitian ini sebesar 79,5% [34].
 9. Penelitian yang dilakukan oleh B. Zhang, X.L. Liang, H.Y. Gao, L.S. Ye dan Y.G. Wang pada tahun 2016 mengenai model analisis regresi logistik, *support vector machine*, dan *back-propagation neural network* dalam diagnosis kanker kolorektal, pada penelitian ini *dataset* yang digunakan berasal dari *Military General Hospital of Beijing PLA*. Adapun metode yang diusulkan adalah *Neural Network* dan SVM, sedangkan hasil yang diperoleh yaitu akurasi untuk NN sebesar 82,5 dan SVM sebesar 75%, sensitivities dari NN sebesar 85% dan SVM sebesar 80%, serta specificities dari NN sebesar 80% dan SVM sebesar 70% [35].
 10. Penelitian yang dilakukan oleh Yiqing Pan, Lei Zhang, Rongrong Zhang, Jing Han, Wenjun Qin, Yong Gu, Jichen Sha, Xiaoyan Xu, Yi Feng, Zhipeng Ren, Jiawen Dai, Ben Huang, Shifang Ren, dan Jianxin Gu pada tahun 2021.

Penelitian ini menggunakan *dataset* yang berasal Shang hai East Hospital, Tongji University, Shanghai, China dengan metode *Random Forest*, J48, *Random Tree*, LMT dan SVM. Adapun penelitian ini memperoleh akurasi klasifikasi 75% [36].

11. Penelitian yang dilakukan oleh Julian Gründner, Hans-Ulrich Prokosch, Michael Stürzl, Roland Croner, Jan Christoph dan Dennis Toddenroth pada tahun 2018 mengenai prediksi hasil klinis pada kanker kolorektal menggunakan metode *machine learning*. Penelitian ini menggunakan *dataset* yang berasal dari Rumah Sakit Universitas Erlangen dengan menggunakan metode *random forest*, *general linear model*, dan *neural network*, dimana hasil akurasi yang diperoleh pada penelitian ini sebesar 71% [37].
12. Penelitian yang dilakukan oleh Xingzhi Yue, Neofytos Dimitriou, Peter D. Caie, David J. Harrison, dan Ognjen Arandjelović pada tahun 2019 mengenai prediksi kanker kolorektal menggunakan *machine learning*. Penelitian ini dilakukan dengan menggunakan metode CNN dan *dataset* yang berasal dari rumah sakit NHS Lothian. Adapun hasil yang diperoleh pada penelitian yaitu Akurasi sebesar 70% dan *F1 Score* sebesar 0.67% [38].
13. Penelitian yang dilakukan oleh Nuntawut Kaoungku, Kittisak Kerdprasop, Nittaya Kerdprasop, Ratiporn Chanklan, dan Keerachart Suksut pada tahun 2019 mengenai klasifikasi citra histologi kanker kolorektal menggunakan *stacked ensembles*. Pada penelitian ini digunakan citra histologi kanker kolorektal yang berasal dari Pusat Medis Universitas Mannheim, Universitas Heidelberg, Mannheim, Jerman. Adapun metode yang diusulkan adalah KNN, *decision tree*, *neural network*, *logistic regression*, SVM, dimana hasil yang diperoleh yaitu KNN sebesar 79%, *decision tree* 36%, *neural network* 86%, *Logistic Regression* 84%, SVM 88% [39].
14. Penelitian yang dilakukan oleh Ole-Johan Skrede, Sepp De Raedt, Andreas Kleppe, Tarjei S Hveem, Knut Liestøl, John Maddison, Hanne A Askautrud, Manohar Pradhan, John Arne Nesheim, Fritz Albrechtsen, Inger Nina Farstad, Enric Domingo, David N Church, Arild Nesbakken, Neil A Shepherd, Ian Tomlinson, Rachel Kerr, Marco Novelli, David J Kerr, dan Håvard E Danielsen mengenai prediksi kanker kolorektal menggunakan teknik *deep*

learning dengan algoritma CNN. Penelitian ini dilakukan dengan menggunakan *dataset* yang berasal dari Rumah Sakit Universitas Akershus, dimana hasil yang didapat pada penelitian ini yaitu akurasi sebesar 76% [40].

15. Penelitian yang dilakukan oleh Neofytos Dimitriou, Ognjen Arandjelović, David J. Harrison dan Peter D. Caie pada tahun 2018 mengenai Kerangka kerja *machine learning* yang berprinsip meningkatkan akurasi prognosis kanker kolorektal stadium II. Penelitian ini dilakukan dengan menggunakan metode KNN dan *dataset* yang berasal dari Rumah sakit Glasgow dan Oxford, dimana hasil penelitian menunjukkan nilai akurasi sebesar 81,1% [41].

Tabel 2.2 Rangkuman Penelitian Terkait

No	Penelitian	Persamaan	Perbedaan
1	Colorectal Histology CSV Multi-classification Accuracy Comparison using Various Machine Learning Models (2019). Metode: CNN, KNN, Logistic Regression, and Random Forest. Hasil: CNN 82,2%, KNN 62,56%, RF 68,72% dan LR 52,6%.	Menggunakan dataset yang berasal dari Pusat Medis Universitas Mannheim, Universitas Heidelberg, Jerman dan salah satu algoritma yang diusulkan sama yaitu CNN.	Menggunakan arsitektur CNN yang berbeda, dimana pada penelitian ini digunakan 3layer dengan parameter lain yang berbeda.
2	Multiple Morphological Constraints-Based Complex Gland Segmentation in Colorectal Cancer Pathology Image Analysis (2016) Metode: CNN Hasil: Akurasi sebesar 86,91%	Menggunakan metode yang sama yaitu CNN	Menggunakan dataset yang berbeda dan jumlah kelas yang lebih sedikit yaitu dua jenis jaringan kanker kolorektal serta arsitektur CNN yang berbeda.
3	Grading of Colorectal Cancer using Histology Images (2016) Metode: SVM Hasil: Akurasi sebesar 81%	Melakukan penelitian mengenai kanker kolorektal dengan menggunakan citra histologi jaringan kanker kolorektal	Menggunakan dataset yang berbeda yaitu Warwick-QU dataset serta metode yang berbeda.

No	Penelitian	Persamaan	Perbedaan
4	<p>Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of Colorectal Cancer (2020)</p> <p>Metode: ResNet-18 dan ResNet-50</p> <p>Hasil: ResNet50 sebesar 80%, sensitivitas 87%, dan spesifisitas 83%.</p>	<p>Melakukan penelitian mengenai kanker kolorektal dengan menggunakan citra histologi jaringan kanker kolorektal</p>	<p>Menggunakan dataset yang berbeda yaitu Warwick-QU dataset serta metode yang berbeda.</p>
5	<p>Multi-class texture analysis in colorectal cancer histology (2016)</p> <p>Metode: SVM</p> <p>Hasil: Akurasi sebesar 87,4%</p>	<p>Menggunakan dataset yang berasal dari Pusat Medis Universitas Mannheim, Universitas Heidelberg, Jerman.</p>	<p>Menggunakan metode yang berbeda.</p>
6	<p>Dimensionality Reduction Strategies for CNN-Based Classification of Histopathological Images (2018)</p> <p>Metode: CNN</p> <p>Hasil: Akurasi sebesar 75%</p>	<p>Menggunakan dataset yang berasal dari Pusat Medis Universitas Mannheim, Universitas Heidelberg, Jerman dan salah satu algoritma yang diusulkan sama yaitu CNN.</p>	<p>Menggunakan arsitektur CNN yang berbeda.</p>
7	<p>Deep Learning with Sampling in Colon Cancer Histology (2019)</p> <p>Metode: CNN</p> <p>Hasil: Akurasi sebesar 84%</p>	<p>Menggunakan metode yang sama yaitu CNN dan Melakukan penelitian dengan menggunakan citra histologi jaringan kanker kolorektal</p>	<p>Menggunakan dataset yang berbeda, dimana dataset yang digunakan adalah The Cancer Genome Atlas (TCGA)</p>

No	Penelitian	Persamaan	Perbedaan
8	<p>Machine Learning Based Classification of Colorectal Cancer Tumour Tissue in Whole-Slide Images (2018)</p> <p>Metode: CNN</p> <p>Hasil: Akurasi sebesar 79,5%</p>	<p>Menggunakan metode yang sama yaitu CNN dan Melakukan penelitian dengan menggunakan citra histologi jaringan kanker kolorektal</p>	<p>Menggunakan dataset yang berbeda, dimana dataset yang digunakan adalah National Center of Pathology, Lithuania</p>
9	<p>Models of logistic regression analysis, support vector machine, and back-propagation neural network based on serum tumor markers in colorectal cancer diagnosis</p> <p>Metode: Neural Network dan SVM</p> <p>Hasil: Akurasi untuk NN sebesar 82,5% dan SVM sebesar 75%, sensitivities dari NN sebesar 85% dan SVM sebesar 80%, serta specificities dari NN sebesar 80% dan SVM sebesar 70%</p>	<p>Melakukan penelitian mengenai kanker kolorektal</p>	<p>Menggunakan dataset yang berbeda, dimana dataset yang digunakan adalah Military General Hospital of Beijing PLA dan metode yang berbeda.</p>
10	<p>Screening and diagnosis of colorectal cancer and advanced adenoma by Bionic Glycome method and machine learning</p> <p>Metode: Random Forest, J48, Random Tree, LMT dan SVM</p> <p>Hasil: Memperoleh akurasi klasifikasi RF 75%</p>	<p>Melakukan penelitian mengenai kanker kolorektal dengan menggunakan citra histologi jaringan kanker kolorektal.</p>	<p>Menggunakan dataset yang berbeda, dimana dataset yang digunakan adalah Shang hai East Hospital, Tongji University, Shanghai, China dan metode yang berbeda.</p>

No	Penelitian	Persamaan	Perbedaan
11	<p>Predicting Clinical Outcomes in Colorectal Cancer Using Machine Learning</p> <p>Metode: Random Forest, General Linear Model, dan Neural Network</p> <p>Hasil: Akurasi RF sebesar 71%</p>	<p>Melakukan penelitian mengenai kanker kolorektal</p>	<p>Menggunakan dataset yang berbeda, dimana dataset yang digunakan adalah Erlangen University Hospital dan metode yang berbeda</p>
12	<p>Colorectal Cancer Outcome Prediction from H&E Whole Slide Images using Machine Learning and Automatically Inferred Phenotype Profiles</p> <p>Metode: CNN</p> <p>Hasil: Akurasi sebesar 70% dan F1 Score sebesar 0.67%</p>	<p>Melakukan penelitian mengenai kanker kolorektal dengan menggunakan citra histologi jaringan kanker kolorektal dan mengusulkan metode yang sama yaitu CNN</p>	<p>Menggunakan dataset yang berbeda, dimana dataset yang digunakan berasal dari NHS Lothian hospitals</p>
13	<p>Colorectal Cancer Histology Image Classification Using Stacked Ensembles</p> <p>Metode: KNN, Decision Tree, Neural Network, Logistic Regression, SVM</p> <p>Hasil: KNN sebesar 79%, DT 36%, NN 86%, LR 84%, SVM 88%</p>	<p>Menggunakan dataset yang berasal dari Pusat Medis Universitas Mannheim, Universitas Heidelberg, Jerman.</p>	<p>Menggunakan metode penelitian yang berbeda, dimana pada penelitian ini diusulkan metode KNN, Decision Tree, Neural Network, Logistic Regression, SVM.</p>
14	<p>Deep learning for prediction of colorectal cancer outcome: a discovery and validation study</p> <p>Metode: CNN</p> <p>Hasil: Akurasi sebesar 76%</p>	<p>Melakukan penelitian mengenai kanker kolorektal dengan menggunakan citra histologi jaringan kanker kolorektal dan mengusulkan metode yang sama yaitu CNN</p>	<p>Menggunakan dataset yang berbeda, dimana dataset yang digunakan berasal dari Akershus University Hospital</p>

No	Penelitian	Persamaan	Perbedaan
15	<p>A principled machine learning framework improves accuracy of stage II colorectal cancer prognosis</p> <p>Metode: KNN</p> <p>Hasil: Akurasi sebesar 81,8%</p>	<p>Melakukan penelitian mengenai kanker kolorektal dengan menggunakan citra histologi jaringan kanker kolorektal</p>	<p>Menggunakan dataset yang berbeda, dimana dataset yang digunakan berasal dari Glasgow and Oxford hospitals dan metode yang berbeda, dimana pada penelitian ini menggunakan metode KNN.</p>

2.3. Tinjauan Objek Penelitian

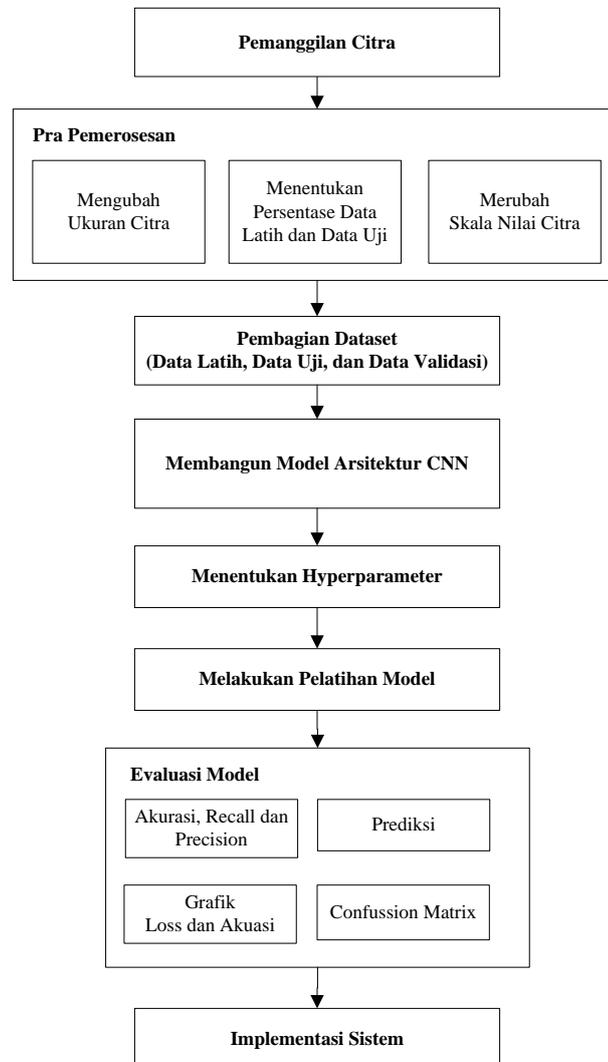
Obyek dari penelitian yang dilakukan menggunakan citra histologi kanker kolorektal yang diperoleh dari situs *kaggle repository* dan berasal dari arsip patologi Institut Patologi, Pusat Medis Universitas Mannheim, Universitas Heidelberg, Mannheim, Jerman. *Dataset* terdiri dari 5.000 citra histologi yang terdiri dari 8 kelas kanker kolorektal diantaranya *empty, tumor, stroma, lympho, debris, mucosa, adipose*, sampai dengan kanker kolorektal yang kompleks dimana masing-masing jumlah citra pada setiap kelas adalah 625. Citra histologi kanker kolorektal ini seluruhnya adalah citra RGB dengan ukuran masing-masing citra sebesar 150 x 150 *pixel*.

BAB III

METODOLOGI PENELITIAN

Bab ini berisi mengenai pemaparan metodologi penelitian yang dilakukan untuk melakukan pengklasifikasian penyakit kanker kolorektal. Tahapan penelitian dilakukan mulai dari pemanggilan citra, pra pemrosesan citra, pembagian dataset menjadi data latih, data uji dan data validasi, membangun arsitektur model yang akan digunakan, penentuan *hyperparameter* model, melakukan pelatihan terhadap model yang diusulkan, melakukan evaluasi model dan yang terakhir melakukan implementasi model menjadi sistem berbasis *website*.

3.1 Tahapan Penelitian



Gambar 3.1 Tahapan Penelitian

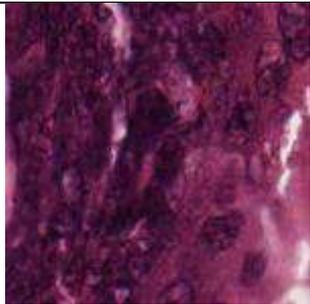
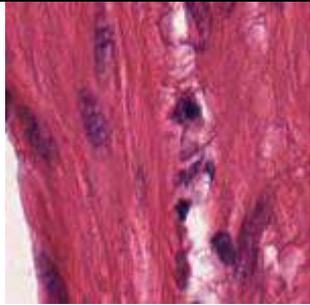
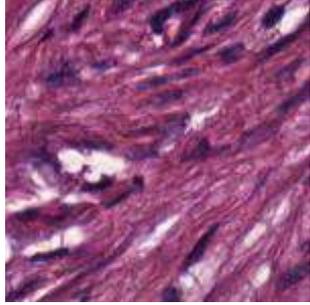
Program Studi Ilmu Komputer (S2) Univeristas Nusa Mandiri

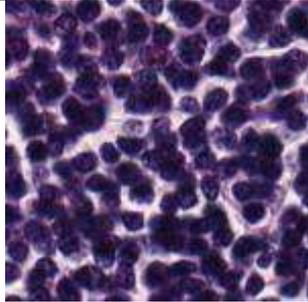
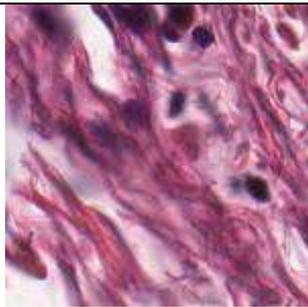
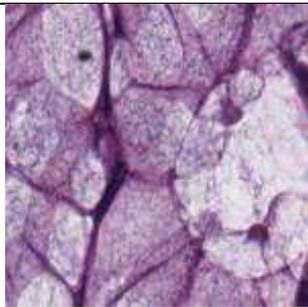
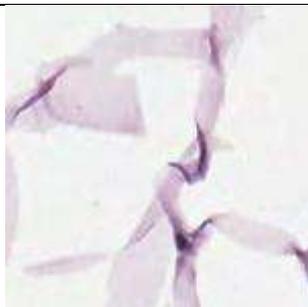
Berdasarkan gambar 3.1 mengenai tahapan penelitian dapat dijelaskan sebagai berikut:

3.1.1 Pemanggilan Citra

Langkah pertama yang dilakukan untuk pemanggilan citra yaitu menentukan direktori penyimpanan *dataset* yang akan digunakan. *Dataset* citra histologi kanker kolorektal memiliki 8 kelas antara lain kelas *empty*, *tumor*, *stroma*, *lympho*, *debris*, *mucosa*, *adipose* dan *complex*. Berikut merupakan *sample* citra setiap kelas *dataset* yang digunakan pada penelitian ini.

Tabel 3.1 Sample Citra Dataset Kanker Kolorektal

No	Contoh Citra	Kelas	Jumlah
1		<i>Tumor</i>	625
2		<i>Stroma</i>	625
3		<i>Complex</i>	625

4		<i>Lympho</i>	625
5		<i>Debris</i>	625
6		<i>Mucosa</i>	625
7		<i>Adipose</i>	625
8		<i>Empty</i>	625
			5.000

Setelah memanggil direktori *dataset* maka langkah selanjutnya adalah memanggil seluruh *library* yang dibutuhkan dalam pembangunan model arsitektur. Diantaranya yaitu *module* OS, *pandas*, *glob*, *cv2*, *numpy*, *matplotlib*, *seaborn*, *sklearn*, *tensorflow*, *keras*.

3.1.2 Pra Pemrosesan

Pada tahapan ini terdiri dari tiga proses tahapan diantaranya merubah ukuran citra menjadi 50x50 piksel, hal ini dilakukan agar citra berukuran kecil sehingga dapat mempercepat proses pengenalan citra pada tahap pemodelan, tahap selanjutnya adalah menentukan persentase data latih dan data uji, yang mana rasio yang digunakan adalah 90:10, artinya 90% digunakan untuk data latih, dan 10% untuk data uji. Kemudian merubah skala nilai citra, hal ini berfungsi untuk merepresentasikan citra menjadi array 3D atau disebut *tensor*, dengan nilai rentangan antara 0-255.

3.1.3 Pembagian Dataset

Pada tahap ini dilakukan pembagian *dataset* menjadi data latih, data uji, dan data validasi, dimana persentase pembagian data latih dan data uji sesuai dengan persentase yang telah ditentukan ditahap sebelumnya, yaitu 90% data latih yakni sebanyak 4450 dan 10% data uji yakni sebanyak 500 sedangkan untuk data validasi diambil sebanyak 50 data dari jumlah data latih.

3.1.4 Pembangunan Model CNN

Pada tahap ini dilakukan percobaan untuk menguji model algoritma yang diusulkan, dengan mempertimbangkan tahapan pemilihan model dan menggunakan arsitekturnya, sehingga dapat disadari bahwa model yang digunakan adalah model yang tepat dari kumpulan data pengujian. Dalam tahap ini dilakukan dengan menentukan jumlah layer arsitektur yang akan digunakan mulai dari menentukan nilai *input layer*, jumlah *convolution layer*, *pooling layer*, *fully connected layer* dan *output layer*, akan tetapi, pada setiap layer diharuskan untuk menentukan beberapa ketentuan seperti pada *convolution layer*, yang mana dalam layer ini ditentukan ukuran *filter*, jumlah kernel, kategori aktivasi, dan *padding*, kemudian pada *pooling*

layer ditentukan kategori *pooling* yang akan digunakan, hal ini bertujuan untuk mereduksi ukuran citra sebanyak mungkin. Tahap selanjutnya adalah menentukan jumlah *dense* pada *fully connected layer*, yang mana nilai ini menandakan jumlah *node* yang harus ada di *hidden layer*, nilainya antara jumlah *input node* dan *output node*. Kemudian menentukan kategori aktivasi yang digunakan, setelah itu menentukan jumlah *dense* pada *output layer*, yang memiliki nilai sesuai dengan jumlah kelas, serta menentukan kategori aktivasi yang berfungsi untuk melakukan perubahan data input menjadi dimensi yang lebih tinggi dengan tujuan agar dapat dilakukan pemotongan batas pemisah (*hyperlane*) sederhana dalam proses klasifikasi. Selain itu, dilakukan pula penambahan nilai *dropout* untuk mencegah terjadinya *overfitting* dan juga mempercepat proses *learning*. Tahapan yang dilakukan dalam pembuatan model CNN dirincikan pada tabel 3.2 sebagai usulan CNN yang dilakukan dalam eksperimen berikut.

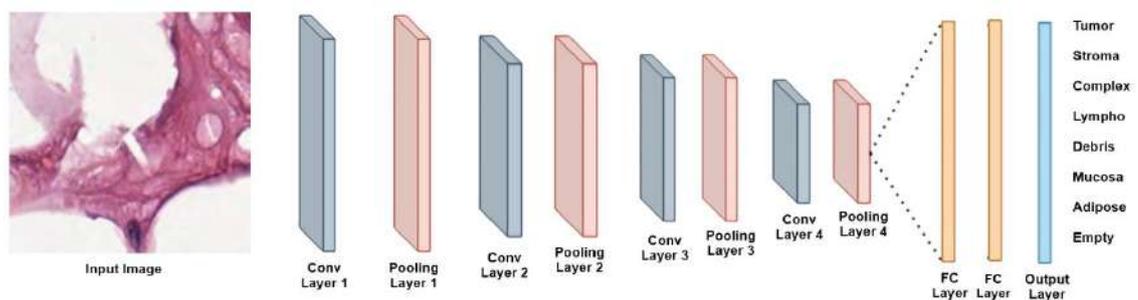
Tabel 3. 2 Arsitektur Usulan CNN

<i>Layer</i>		<i>Size dan Filters</i>	<i>Kernel Size</i>	<i>Activation</i>	<i>Pooling size</i>
1	<i>Image</i>	50 x 50 x 3	-	-	-
	<i>Covolution</i>	16x16	2x2	<i>relu</i>	-
	<i>Maxpooling</i>	-	-	-	2x2
2	<i>Covolution</i>	32x32	2x2	<i>relu</i>	-
	<i>Maxpooling</i>	-	-	-	2x2
3	<i>Covolution</i>	64x64	2x2	<i>relu</i>	-
	<i>Maxpooling</i>	-	-	-	2x2
4	<i>Covolution</i>	128x128	2x2	<i>relu</i>	-
	<i>Maxpooling</i>	-	-	-	2x2
<i>Flatten</i>		-	-	-	-
<i>Dropout</i>		0,2			
<i>Dense</i>		200		<i>relu</i>	
<i>Dropout</i>		0,3			
<i>Dense</i>		500		<i>relu</i>	
<i>Dropout</i>		0,4			
<i>Dense</i>		8		<i>softmax</i>	

Penelitian ini mengusulkan empat layer konvolusi dengan input citra sebesar 50x50 sesuai dengan ukuran citra yang telah diinisiasikan sebelumnya. Operasi konvolusi pada layer pertama menggunakan kernel ukuran 2x2 dengan filter sebanyak 16 buah sehingga menghasilkan *output feature map* berukuran 16x16 sebanyak 16. kemudian seluruh *feature map* diaktivasi dengan memetakan setiap nilai *pada feature map* terhadap fungsi aktivasi ReLu. Setelah itu masuk pada bagian *pooling*, pada penelitian ini jenis *pooling* yang digunakan adalah *maximum pooling* dengan ukuran 2x2, sehingga menghasilkan 16 buah feature map.

Selanjutnya konvolusi pada bagian kedua dilakukan dengan kernel 2x2 dengan *filter* 32x32 dengan *activation* ReLu dan *maxpooling* 2x2, konvolusi ketiga dilakukan dengan kernel 2x2 dengan *filter* 64x64 dengan *activation* ReLu dan *maxpooling* 2x2 dan layer ke empat dilakukan konvolusi dengan kernel 2x2 dan *filter* berukuran 128x128.

Kemudian pada proses konvolusi menambahkan fungsi *zero padding* dimana *zero padding* merupakan parameter piksel bernilai nol yang ditambahkan pada setiap sisi input. Setelah melalui proses konvolusi hasil akhir dari *max pooling* akan diubah kedalam bentuk *vector* dua dimensi. Pada *hidden layer* menggunakan *dense* dengan jumlah *vector* 200, nilai *dropout* 0,2 serta *dense* dengan jumlah *vector* 500 dan nilai *dropout* 0,3 serta menggunakan aktivasi fungsi ReLU. Pada *output layer* terdapat *dense* sebesar 8 sesuai dengan jumlah kelas yang ada pada *dataset* dan menambahkan nilai *dropout* sebesar 0,4 serta fungsi aktivasi *softmax*. Fungsi aktivasi ini bertujuan untuk mengklasifikasi kedalam banyak kelas.



Gambar 3. 1 Model CNN

Setelah melalui *forward phase* maka akan dilakukan *backward* pada model CNN untuk menghitung *gradient* setiap parameter yang akan dioptimasi berupa

weight dan bias. Kemudian parameter tersebut diperbaharui dengan parameter baru berdasarkan algoritma optimasi *RMSProp*, dan proses ini disebut sebagai *backward phase*. Adapun proses *backpropagation* merupakan proses yang dilakukan secara keseluruhan untuk memproses *forward phase* hingga *backward phase*. Sedangkan proses *backpropagation* yang dilakukan pada keseluruhan data dinamakan *epoch*. Pada penelitian ini proses dilakukan sebanyak 100 *epoch* dengan ukuran *batch size* sebesar 50.

3.1.5. Penentuan *Hyperparameter*

Tahap selanjutnya adalah menentukan *hyperparameter* untuk model arsitektur yang dibangun mulai dari jenis optimasi, jenis *loss* probabilistik serta pengukuran evaluasi model yang akan digunakan, dimana pada penelitian ini diusulkan penggunaan jenis optimasi *RMSProp*, *categorical loss cross entropy loss* dan pengukuran evaluasi dengan menggunakan parameter evaluasi.

3.1.6. Pelatihan Model

Tahap keenam merupakan tahap pelatihan model dari arsitektur yang telah dibangun, dalam tahap ini menentukan jumlah *epoch* yang akan digunakan, dimana jumlah *epoch* adalah *hyperparameter* yang menentukan berapa kali algoritma pembelajaran akan bekerja melalui seluruh *dataset* pelatihan, serta menentukan jumlah *batch size* yang merupakan jumlah sampel yang diproses sebelum model di *update*. Adapun jumlah *epoch* yang digunakan sebanyak 100 dan *batch size* yang digunakan sebesar 50.

3.1.7. Evaluasi Model

Tahap evaluasi model berfungsi untuk melihat performa model yang telah dibangun, tahap ini dapat dilihat dari nilai akurasi yang diperoleh, prediksi untuk menguji sampel citra berdasarkan model, melihat grafik yang dihasilkan dari iterasi pelatihan model, serta tabel *confusion matrix*.

3.1.8. Implementasi Sistem

Tahap terakhir merupakan tahap implementasi dari model arsitektur yang telah dibangun menjadi sebuah sistem prediksi, dimana sistem ini akan dibangun dengan menggunakan *flask web development* dengan menggunakan bahasa pemrograman *Python*.



Gambar 3. 3 Rancangan Antarmuka Aplikasi

Sistem akan dibuat dimana *user* hanya perlu memasukan citra salah satu jenis jaringan kanker kolorektal kemudian menekan tombol *process* dan sistem akan memproses citra tersebut berdasarkan alur model yang telah dibuat dengan menggunakan metode CNN. Setelah itu sistem akan menampilkan kategori jaringan penyakit kanker kolorektal yang telah dipilih.

3.2. Instrumen Penelitian

Instrumen yang digunakan pada penelitian ini menggunakan *windows 10* sebagai *operating system*, dan spesifikasi perangkat komputer prosesor *Intel Core i5-1035G1 CPU @ 1.00 Ghz* dengan RAM sebesar 8.00 GB DDR4 serta VGA Intel UHD Graphics. Penelitian dilakukan dengan menggunakan *Google Colab* dimana Bahasa pemrograman yang digunakan adalah *Python* versi 3 serta implementasi sistem berbasis web menggunakan *Python Flask*.

3.3. Metode Pengumpulan Data

Penelitian ini menggunakan data sekunder yaitu data yang telah dikumpulkan dari peneliti sebelumnya [31] dimana Sampel histologis merupakan citra yang sepenuhnya dianonimkan dari adenokarsinoma kolorektal manusia (tumor primer) yang tertanam parafin tetap formalin dan berasal dari arsip patologi (Institute of Pathology, University Medical Center Mannheim, Heidelberg University, Mannheim, Jerman). Dataset tersebut memiliki 5.000 citra histologi dengan 8 jenis jaringan kanker kolorektal diantaranya:

1. Tumor atau tumor maligna merupakan tumor yang muncul dari jaringan epitel kolon atau rektum.
2. *Stroma* merupakan jenis kanker yang menyerang sel-sel khusus dinding kolon yang disebut *interstitial cajal*.
3. *Complex* atau Stroma kompleks merupakan jenis kanker kolorektal yang menyerang dinding kolon akan tetapi jenis kanker ini sudah memasuki tingkat keparahan yang cukup tinggi, stroma kompleks mengandung sel tumor tunggal dan atau sedikit sel imun.
4. *Lympho* merupakan jenis kanker kolorektal yang telah menembus dinding kolon juga dapat menembus darah atau kelenjar getah bening (*lymph vessels*)
5. *Debris* merupakan puing-puing sisa sel tumor (termasuk nekrosis, perdarahan dan lendir).
6. *Mucosa* merupakan sel-sel kanker masih sebatas pada lapisan dinding usus besar (lapisan mukosa). Jenis kanker ini termasuk kategori kanker kolorektal stadium 1.
7. *Adipose* merupakan gejala yang disebabkan karena jaringan adiposa menghasilkan banyak hormon dan sitokin pro-inflamasi diantaranya yaitu, interleukin 6, *Tumor Necrotic Factor* ($TNF-\alpha$), leptin dan adiponektin yang memberikan lingkungan mikro inflamasi yang dibutuhkan sebagai tempat hidup sel kanker.
8. *Empty* atau tidak terdeteksi tumor kanker kolorektal.

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

Bab ini berisi mengenai pemaparan hasil penelitian yang telah dilakukan mulai dari pembahasan mengenai beberapa skenario yang dirancang, hasil yang didapatkan pada setiap pengujian yang dilakukan serta analisa dari setiap pengujian tersebut.

4.1. Skenario Uji Coba

Uji coba pada penelitian ini dilakukan pada 5.000 citra histologi jaringan penyakit kanker kolorektal yang terdiri dari 8 kelas, dimana pada setiap kelas data terdiri dari 625 citra. Pada bagian ini berisi skema uji coba untuk mengukur tingkat keberhasilan metode yang diusulkan, dimana pengujian dilakukan dengan melakukan beberapa percobaan diantaranya uji coba terhadap metode yang diusulkan, uji coba dengan melakukan perbandingan penerapan jumlah layer konvolusi, jumlah *epoch*, persentase rasio data uji dan data latih yang digunakan, perbandingan jenis *optimizer* serta uji coba terhadap metode usulan dengan metode *transfer learning* yaitu *Mobilenet*, *Densenet*, *VGG16*, dan *Restnet50*.

4.1.1. Hasil Uji Coba Metode Usulan

Hasil uji coba dilakukan dengan menggunakan persentase data uji sebesar 90% dan data latih sebesar 10%, kemudian ukuran citra yang digunakan sebesar 50x50 *pixels*, hal ini dilakukan agar citra berukuran kecil sehingga dapat mempercepat proses pengenalan citra pada tahap pemodelan. Citra yang digunakan pada penelitian ini seluruhnya menggunakan citra RGB. Adapun arsitektur CNN yang diusulkan menggunakan 4 layer konvolusi dimana pada layer pertama jumlah *filter* yang digunakan sebesar 16x16, pada layer konvolusi yang kedua digunakan jumlah *filter* sebesar 32x32, kemudian pada layer ketiga digunakan jumlah *filter* sebesar 64x64 serta pada layer keempat digunakan jumlah filter sebesar 128x128. Adapun jumlah kernel yang digunakan pada setiap layer yaitu sebesar 2x2 dan jenis aktivasi yang dipilih yaitu Relu serta jenis *pooling* yang digunakan pada setiap layer

adalah *maxpooling* dengan ukuran stride sebesar 2x2. Pada metode usulan ini diterapkan 2 *fully connected layer* dimana jumlah *dense* yang digunakan pada *fully connected layer* pertama sebesar 200 dan pada *fully connected layer* kedua digunakan *dense* sebesar 500 serta pada setiap layer, dilakukan pula penambahan nilai *dropout* untuk mencegah terjadinya *overfitting* dan juga mempercepat proses *learning*. Nilai *dropout* yang digunakan pada layer *fully connected* pertama sebesar 0.2 sedangkan pada layer *fully connected* kedua sebesar 0.3 serta pada *output layer* digunakan jenis aktivasi *softmax*. Total *param* yang didapat dari arsitektur yang diusulkan sebanyak 378,548. Pelatihan model dilakukan dengan menggunakan 100 *epoch*, ukuran *batch size* sebesar 50, dan jenis optimasi yang dipilih yaitu RMSProp serta jenis *loss* probabilistik yang digunakan yaitu *categorical_crossentropy*. Nilai akurasi yang dihasilkan dari model yang diusulkan yaitu sebesar 0.8960000276565552. Berikut merupakan model dari arsitektur yang dibangun.

```
Model: "sequential_1"
```

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 50, 50, 16)	208
max_pooling2d_4 (MaxPooling2D)	(None, 25, 25, 16)	0
conv2d_5 (Conv2D)	(None, 25, 25, 32)	2080
max_pooling2d_5 (MaxPooling2D)	(None, 12, 12, 32)	0
conv2d_6 (Conv2D)	(None, 12, 12, 64)	8256
max_pooling2d_6 (MaxPooling2D)	(None, 6, 6, 64)	0
conv2d_7 (Conv2D)	(None, 6, 6, 128)	32896
max_pooling2d_7 (MaxPooling2D)	(None, 3, 3, 128)	0
flatten_1 (Flatten)	(None, 1152)	0
dropout_3 (Dropout)	(None, 1152)	0
dense_3 (Dense)	(None, 200)	230600
dropout_4 (Dropout)	(None, 200)	0
dense_4 (Dense)	(None, 500)	100500
dropout_5 (Dropout)	(None, 500)	0
dense_5 (Dense)	(None, 8)	4008

```

Total params: 378,548
Trainable params: 378,548
Non-trainable params: 0

```

Gambar 4.1. Usulan Model Arsitektur CNN

4.1.2. Hasil Uji Coba Jumlah Layer

Skenario kedua dilakukan dengan menguji jumlah lapisan konvolusi yang digunakan pada model pengujian. Hal ini dilakukan untuk membandingkan nilai akurasi pada saat menggunakan 1 layer, 2 layer, 3 layer, 4 layer, dan 5 layer konvolusi secara berurutan untuk arsitektur CNN yang telah diusulkan, dalam pengujian ini ukuran kernel, ukuran filter, penggunaan *padding*, aktivasi, jumlah *dense* serta *dropout* pada lapisan *fully connected* bahkan jenis optimasi, jenis *loss* probabilistik, jumlah epoch serta jumlah *batch size* tetap menggunakan parameter yang sama dengan model yang di usulkan. Hal ini bertujuan untuk mengukur tingkat keberhasilan metode yang diusulkan. Berikut merupakan masing-masing arsitektur berdasarkan skenario uji coba layer yang digunakan.

Tabel 4. 1 Arsitektur CNN 1 Layer Konvolusi

<i>Layer</i>		<i>Size dan Filters</i>	<i>Kernel Size</i>	<i>Activation</i>	<i>Pooling size</i>
1	<i>Image</i>	50 x 50 x 3	-	-	-
	<i>Covolution</i>	16x16	2x2	<i>relu</i>	-
	<i>Maxpooling</i>	-	-	-	2x2
<i>Flatten</i>		-	-	-	-
<i>Dropout</i>		0,2			
<i>Dense</i>		200		<i>relu</i>	
<i>Dropout</i>		0,3			
<i>Dense</i>		500		<i>relu</i>	
<i>Dropout</i>		0,4			
<i>Dense</i>		8		<i>softmax</i>	

Tabel 4. 2 Arsitektur CNN 2 Layer Konvolusi

<i>Layer</i>		<i>Size dan Filters</i>	<i>Kernel Size</i>	<i>Activation</i>	<i>Pooling size</i>
1	<i>Image</i>	50 x 50 x 3	-	-	-
	<i>Covolution</i>	16x16	2x2	<i>relu</i>	-
	<i>Maxpooling</i>	-	-	-	2x2
2	<i>Covolution</i>	32x32	2x2	<i>relu</i>	-
	<i>Maxpooling</i>	-	-	-	2x2
<i>Flatten</i>		-	-	-	-
<i>Dropout</i>		0,2			
<i>Dense</i>		200		<i>relu</i>	
<i>Dropout</i>		0,3			
<i>Dense</i>		500		<i>relu</i>	
<i>Dropout</i>		0,4			
<i>Dense</i>		8		<i>softmax</i>	

Tabel 4. 3 Arsitektur CNN 3 Layer Konvolusi

Layer		Size dan Filters	Kernel Size	Activation	Pooling size
1	Image	50 x 50 x 3	-	-	-
	Covolution	16x16	2x2	relu	-
	Maxpooling	-	-	-	2x2
2	Covolution	32x32	2x2	relu	-
	Maxpooling	-	-	-	2x2
3	Covolution	64x64	2x2	relu	-
	Maxpooling	-	-	-	2x2
Flatten		-	-	-	-
Dropout		0,2			
Dense		200		relu	
Dropout		0,3			
Dense		500		relu	
Dropout		0,4			
Dense		8		softmax	

Tabel 4. 4 Arsitektur CNN 4 Layer Konvolusi

Layer		Size dan Filters	Kernel Size	Activation	Pooling size
1	Image	50 x 50 x 3	-	-	-
	Covolution	16x16	2x2	relu	-
	Maxpooling	-	-	-	2x2
2	Covolution	32x32	2x2	relu	-
	Maxpooling	-	-	-	2x2
3	Covolution	64x64	2x2	relu	-
	Maxpooling	-	-	-	2x2
4	Covolution	128x128	2x2	relu	-
	Maxpooling	-	-	-	2x2
Flatten		-	-	-	-
Dropout		0,2			
Dense		200		relu	
Dropout		0,3			
Dense		500		relu	
Dropout		0,4			
Dense		8		softmax	

Tabel 4. 5 Arsitektur CNN 5 Layer Konvolusi

Layer		Size dan Filters	Kernel Size	Activation	Pooling size
1	Image	50 x 50 x 3	-	-	-
	Covolution	16x16	2x2	relu	-
	Maxpooling	-	-	-	2x2

2	<i>Covolution</i>	32x32	2x2	<i>relu</i>	-
	<i>Maxpooling</i>	-	-	-	2x2
3	<i>Covolution</i>	64x64	2x2	<i>relu</i>	-
	<i>Maxpooling</i>	-	-	-	2x2
4	<i>Covolution</i>	128x128	2x2	<i>relu</i>	-
	<i>Maxpooling</i>	-	-	-	2x2
5	<i>Covolution</i>	256x256	2x2	<i>relu</i>	-
	<i>Maxpooling</i>	-	-	-	2x2
	<i>Flatten</i>	-	-	-	-
	<i>Dropout</i>	0,2			
	<i>Dense</i>	200		<i>relu</i>	
	<i>Dropout</i>	0,3			
	<i>Dense</i>	500		<i>relu</i>	
	<i>Dropout</i>	0,4			
	<i>Dense</i>	8		<i>softmax</i>	

Uji coba penggunaan jumlah layer konvolusi dilakukan dengan tetap menggunakan parameter yang sama, dimana nilai *filter* yang diujikan pada penelitian ini akan bertambah sejumlah dua kali lipat dari nilai filter sebelumnya yaitu pada uji coba penggunaan 1 layer konvolusi jumlah filter yang digunakan sebesar 16x16, kemudian pengujian dengan menggunakan 2 layer konvolusi jumlah filter yang digunakan pada layer pertama sebesar 16x16 dan pada layer kedua sebesar 32x32, pengujian dengan menggunakan 3 layer konvolusi jumlah filter yang digunakan pada layer pertama sebesar 16x16 kemudian layer kedua sebesar 32x32 dan pada layer ketiga sebesar 64x64, pengujian dengan menggunakan 4 layer konvolusi jumlah filter yang digunakan pada layer pertama sebesar 16x16 kemudian layer kedua sebesar 32x32, setelah itu pada layer ketiga sebesar 64x64 dan pada layer keempat sebesar 128x128.

Uji coba skenario layer konvolusi juga dilakukan dengan menguji 5 layer konvolusi, dimana jumlah filter yang digunakan pada layer pertama sebesar 16x16 kemudian layer kedua sebesar 32x32, setelah itu pada layer ketiga sebesar 64x64 dan pada layer keempat sebesar 128x128 serta pada layer kelima sebesar 256x256.

Adapun jumlah kernel yang digunakan pada uji coba skenario penggunaan layer konvolusi sebesar 2x2, dengan jenis aktivasi yang digunakan adalah *relu*, serta jenis *pooling* yang digunakan pada setiap layer adalah *maxpooling* dengan ukuran stride sebesar 2x2. Pada uji coba layer konvolusi diterapkan pula 2 *fully connected layer* dimana jumlah *dense* yang digunakan pada *fully connected layer* pertama

sebesar 200 dan pada *fully connected layer* kedua digunakan *dense* sebesar 500 serta pada setiap layer, dilakukan pula penambahan nilai *dropout*. Nilai *dropout* pada layer *fully connected* pertama sebesar 0,2 sedangkan pada layer *fully connected* kedua sebesar 0,3 serta pada *output layer* digunakan jenis aktivasi *softmax*. Pelatihan model dilakukan dengan menggunakan 100 *epoch*, ukuran *batch size* sebesar 50, dan jenis optimasi yang dipilih yaitu RMSProp serta jenis *loss* probabilistik yang digunakan yaitu *categorical_crossentropy*. Berikut merupakan hasil pengujian terhadap model dengan jumlah layer konvolusi yang berbeda.

Tabel 4.6. Nilai Akurasi Berdasarkan Jumlah Layer Konvolusi

No	Jumlah Layer	Akurasi
1	1	0.777999997138977
2	2	0.8299999833106995
3	3	0.8460000157356262
4	4	0.8960000276565552
5	5	0.8399999737739563

Berdasarkan Tabel 4.6 nilai akurasi terbesar dihasilkan pada pengujian 4 layer dengan nilai akurasi sebesar 89,60% sesuai dengan model arsitektur CNN yang diusulkan.

4.1.3. Hasil Uji Coba Jumlah *Epoch*

Skenario ketiga dilakukan dengan menguji jumlah *epoch* yang digunakan pada model pengujian. Hal ini dilakukan untuk membandingkan nilai akurasi pada saat menggunakan 100 *epoch*, 50 *epoch* dan 10 *epoch* secara berurutan untuk arsitektur CNN yang telah diusulkan, dalam pengujian ini ukuran kernel, ukuran filter, penggunaan *padding*, aktivasi, jumlah *dense* serta *dropout* pada lapisan *fully connected* bahkan jenis optimasi, jenis *loss* probabilistik, jumlah *epoch* serta jumlah *batch size* tetap menggunakan parameter yang sama dengan model yang diusulkan. Hal ini bertujuan untuk mengukur tingkat keberhasilan metode yang diusulkan. Berikut merupakan hasil pengujian terhadap model dengan jumlah *epoch* yang berbeda.

Tabel 4.7. Nilai Akurasi Berdasarkan Jumlah *Epoch*

No	Jumlah <i>Epoch</i>	Akurasi
1	100	0.8960000276565552
2	50	0.8640000224113464
3	10	0.7639999985694885

Berdasarkan Tabel 4.7 nilai akurasi terbesar dihasilkan pada pengujian 100 *epoch* dengan nilai akurasi sebesar 89,60%. Adapun jumlah *epoch* 50 menghasilkan nilai akurasi sebesar 86,40% sedangkan jumlah *epoch* 10 tidak memberikan nilai akurasi yang baik untuk penelitian ini.

4.1.4. Hasil Uji Coba Persentase Rasio Data Latih dan Data Uji

Skenario keempat dilakukan dengan menguji jumlah persentase data latih dan data uji yang digunakan pada model pengujian. Hal ini dilakukan untuk membandingkan nilai akurasi pada saat menggunakan rasio 60:40, 70:30, 80:20 dan 90:10 secara berurutan untuk arsitektur CNN yang telah diusulkan, dalam pengujian ini ukuran kernel, ukuran *filter*, penggunaan *padding*, aktivasi, jumlah *dense* serta *dropout* pada lapisan *fully connected* bahkan jenis optimasi, jenis *loss* probabilistik, jumlah *epoch* serta jumlah *batch size* tetap menggunakan parameter yang sama dengan model yang di usulkan. Hal ini bertujuan untuk mengukur tingkat keberhasilan metode yang diusulkan. Berikut merupakan hasil pengujian terhadap model dengan persentase data latih dan data uji yang berbeda.

Tabel 4.8. Nilai Akurasi Berdasarkan Persentase Data Uji dan Data Latih

No	Rasio	Akurasi
1	60:40	0.8420000076293945
2	70:30	0.8640000224113464
3	80:20	0.8700000047683716
4	90:10	0.8960000276565552

Berdasarkan Tabel 4.8 nilai akurasi terbesar dihasilkan pada pengujian dengan menggunakan persentase data latih sebesar 90% dan data uji sebesar 10%

dengan nilai akurasi sebesar 89,60% sesuai dengan persentase dataset yang diusulkan.

4.1.5. Hasil Uji Coba Jenis Optimasi

Skenario uji kelima dilakukan dengan menguji jenis optimasi yang digunakan pada model pengujian. Hal ini dilakukan untuk membandingkan nilai akurasi pada saat menggunakan jenis optimasi RMSProp, Adam, SGD, Adadelta dan Adagrad secara berurutan untuk arsitektur CNN yang telah diusulkan, dalam pengujian ini ukuran kernel, ukuran *filter*, penggunaan *padding*, aktivasi, jumlah *dense* serta *dropout* pada lapisan *fully connected* bahkan jenis optimasi, jenis *loss* probabilistik, jumlah *epoch* serta jumlah batch size tetap menggunakan parameter yang sama dengan model yang di usulkan. Hal ini bertujuan untuk mengukur tingkat keberhasilan metode yang diusulkan. Berikut merupakan hasil pengujian terhadap model dengan persentase data latih dan data uji yang berbeda.

Tabel 4.9. Nilai Akurasi Berdasarkan Jenis Optimasi

No	Optimasi	Akurasi
1	RMSProp	0.8960000276565552
2	Adam	0.8659999966621399
3	SGD	0.4259999990463257
4	Adagrad	0.8119999766349792
5	Adadelta	0.4259999990463257

Berdasarkan Tabel 4.9 nilai akurasi terbesar dihasilkan pada pengujian dengan menggunakan optimasi RMSProp yaitu optimasi yang digunakan pada metode usulan dengan nilai akurasi sebesar 89,60%. Adapun optimasi lainnya seperti Adam cukup baik digunakan pada penelitian ini karena menghasilkan nilai akurasi sebesar 86,59% sedangkan optimasi Adadelta dan SGD tidak memberikan nilai akurasi yang baik untuk penelitian ini.

4.1.6. Hasil Uji Metode *Transfer Learning*

Skenario uji coba keenam dilakukan setelah mendapatkan klasifikasi terbaik pada uji coba yang telah dilakukan sebelumnya. Pada tahapan uji coba ini dilakukan dengan menguji metode transfer learning dengan metode yang diusulkan yang bertujuan untuk mengukur performa metode yang dibandingkan. Adapun Metode arsitektur *transfer learning* yang dipilih antara lain *Resnet50*, *VGG16*, *Densenet121* dan *Mobilenet*. Berikut merupakan hasil dari eksperimen yang dilakukan:

Tabel 4.10. Nilai Akurasi Transfer Learning

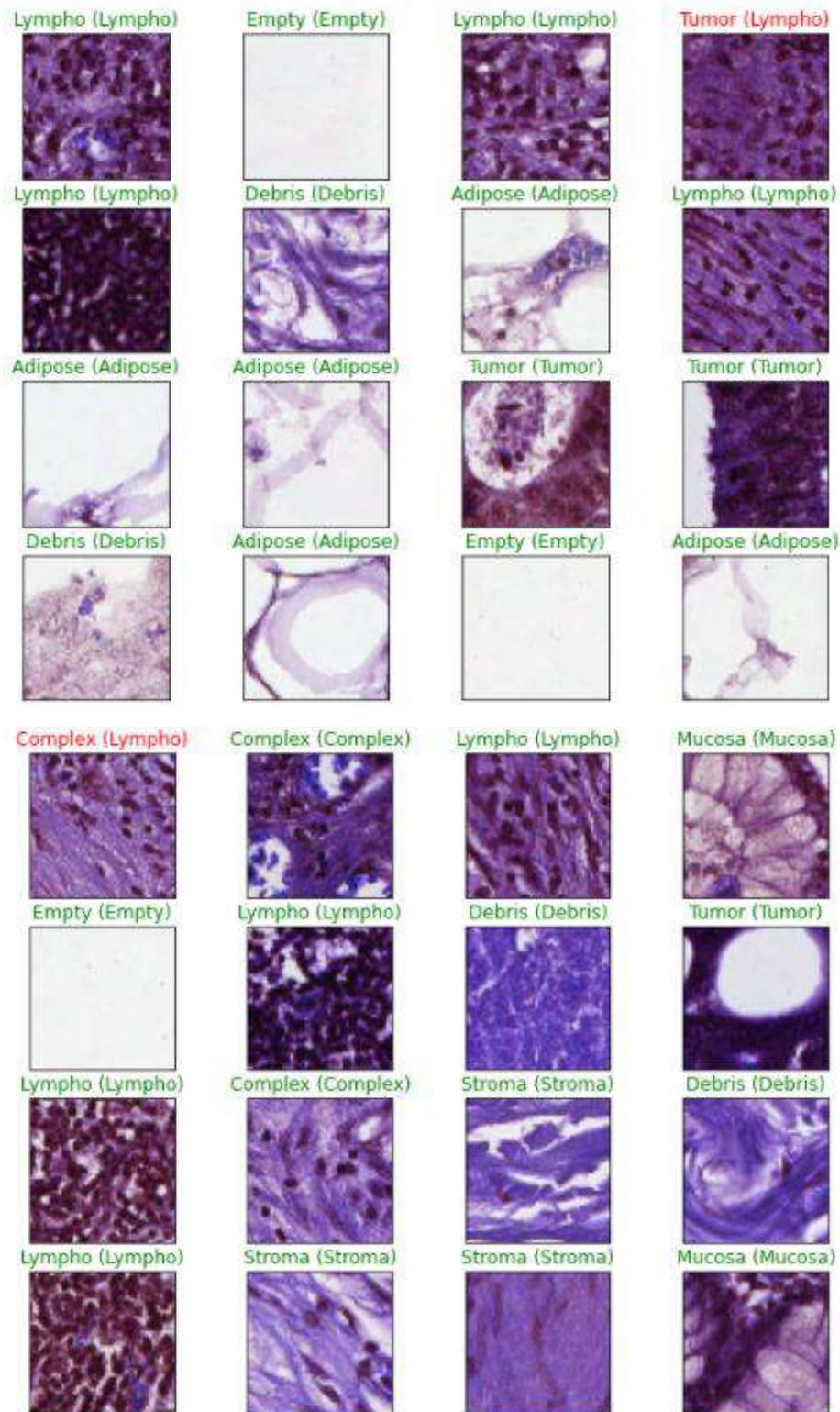
No	TL	Akurasi
1	Mobilenet	0.3540000021457672
2	Restnet 50	0.7440000176429749
3	VGG16	0.8560000061988831
4	Densenet 50	0.8100000023841858
5	Metode Usulan	0.8960000276565552

Berdasarkan Tabel 4.10 nilai akurasi terbesar dihasilkan pada pengujian tetap dengan menggunakan metode usulan CNN *non transfer learning* dengan nilai akurasi sebesar 89,60%. Adapun metode *transfer learning* lainnya seperti arsitektur VGG16 cukup baik digunakan pada penelitian ini karena menghasilkan nilai akurasi sebesar 85,60% sedangkan metode *transfer learning* dengan menggunakan arsitektur *Mobilenet* tidak memberikan nilai akurasi yang baik untuk penelitian ini.

4.2. Evaluasi Hasil Klasifikasi

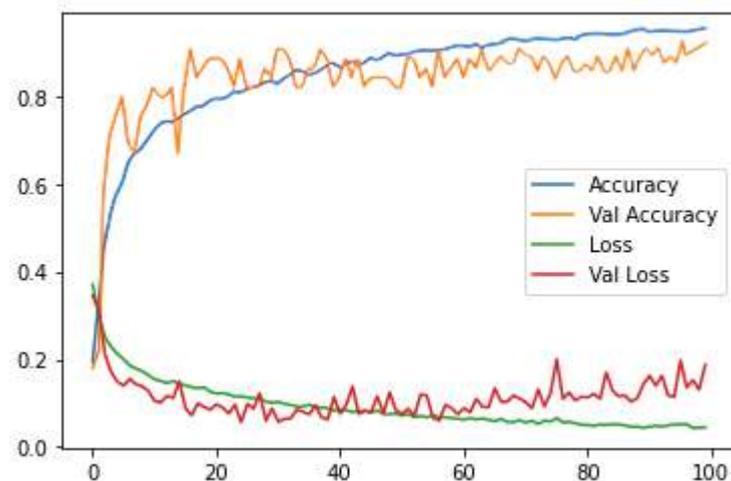
Hasil penelitian yang dilakukan dengan beberapa skenario uji yang telah dilakukan mulai dengan membandingkan arsitektur jumlah layer konvolusi yang digunakan, jumlah persentase rasio dataset yang digunakan, jenis optimasi yang digunakan hingga membandingkan metode usulan CNN *non transfer learning* dengan metode *transfer learning* terlihat bahwa hasil terbaik didapat dengan menggunakan metode CNN *non transfer learning*, dengan jumlah layer konvolusi sebanyak 4 layer, dan jenis optimasi RMSProp serta persentasi dataset yang digunakan sebesar 80:20 memberikan nilai akurasi sebesar 90,60%. Metode CNN

yang diusulkan dapat memprediksi citra jenis jaringan kanker kolorektal dengan baik seperti pada gambar 4.2 berikut.



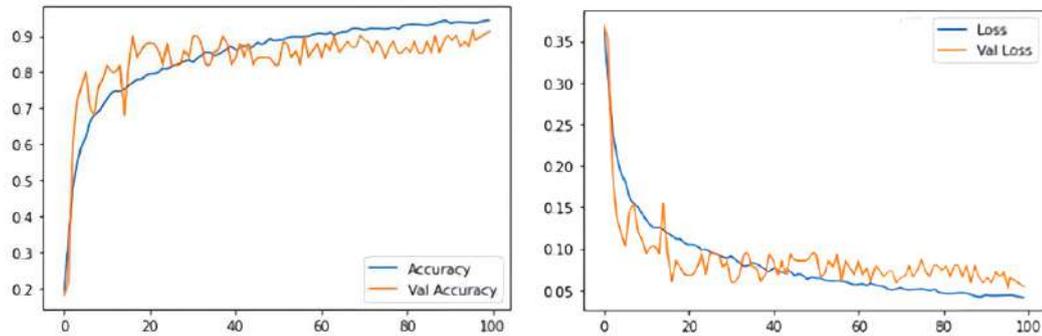
Gambar 4.2 Hasil Prediksi Citra Jenis Jaringan Kanker Kolorektal

Pada Gambar 4.2 menampilkan sebagian prediksi yang dihasilkan oleh model CNN yang diusulkan pada penelitian ini. Berdasarkan gambar tersebut diberikan keterangan jenis jaringan kanker kolorektal dimana keterangan berwarna hijau menunjukkan bahwa prediksi yang dihasilkan bernilai benar seperti pada citra baris kedua dan kolom kedua dimana citra tersebut merupakan jenis jaringan debris dan model terbukti memprediksi bahwa citra tersebut adalah benar jenis jaringan debris sedangkan keterangan yang ditunjukkan dengan warna merah bernilai salah seperti pada baris pertama kolom keempat dimana model memprediksi bahwa citra tersebut adalah Tumor sedangkan jenis jaringan yang sebenarnya adalah *lympho*. Akan tetapi secara keseluruhan model yang diusulkan memprediksi citra dengan benar terbukti dengan banyaknya prediksi yang bernilai benar dibandingkan dengan prediksi yang bernilai salah. Selain itu, hasil dari model yang diusulkan menghasilkan grafik *accuracy*, *val accuracy*, *loss* dan *val loss*.



Gambar 4. 3 Grafik Model Usulan

Pada Gambar 4.3. menampilkan grafik mengenai nilai akurasi yang dihasilkan pada setiap *epoch*. Berdasarkan grafik tersebut dapat disimpulkan bahwa nilai akurasi setiap *epoch* pada data latih rata-rata mengalami kenaikan yang cukup stabil mulai dari *epoch* pertama dengan nilai akurasi sebesar 0.1947 dan terus bertambah sampai dengan *epoch* terakhir dimana pada *epoch* ke 90 menghasilkan nilai akurasi yang tinggi sebesar 0.9354, artinya semakin besar nilai akurasi yang dihasilkan maka semakin baik model yang diusulkan. Sedangkan untuk data validasi mengalami kenaikan maupun penurunan pada *epoch* tertentu.



Gambar 4. 4 Hasil Grafik Accuracy dan Loss

Adapun nilai *loss* pada data uji yang dihasilkan dalam penelitian ini menunjukkan semakin mengalami penurunan dimana nilai *loss* pertama sebesar 0.3705 sampai pada *epoch* ke 100 bernilai 0.0446, nilai *loss* yang semakin kecil menunjukkan tingkat kesalahan yang rendah, akan tetapi nilai *loss* pada data validasi mengalami kenaikan maupun penurunan pada *epoch* tertentu.

Hasil evaluasi lainnya ditunjukkan dengan tabel confusion matrix, dimana pada table ini menunjukkan jumlah prediksi yang bernilai benar atau *true positive* (TP) dan bernilai salah atau *true negative* (TN) maupun hasil prediksi yang bernilai benar akan tetapi sebenarnya adalah salah atau *false positive* maupun menunjukkan hasil prediksi yang bernilai salah akan tetapi seharusnya adalah benar atau *false negative* (FN). Berikut merupakan tabel *confusion matrix* yang dihasilkan pada penelitian ini.

Tabel 4.11. Confusion Matrix

	Prediksi								
		Tumor (0)	Stroma (1)	Complex (2)	Lympho (3)	Debris (4)	Mucosa (5)	Adipose (6)	Empty (7)
Aktual	Tumor (0)	73	0	0	1	0	1	0	0
	Stroma (1)	1	42	4	0	0	0	0	0
	Complex (2)	4	5	41	7	0	1	0	0
	Lympho (3)	2	0	0	53	0	0	0	0
	Debris (4)	0	11	3	0	57	2	1	0

	Mucosa (5)	2	0	0	1	2	59	1	0
	Adipose (6)	0	0	0	0	1	0	61	2
	Empty (7)	0	0	0	0	0	0	0	62

Berdasarkan hasil *confusion matrix* bernilai *true label* dengan *predicted label* yaitu tumor sebanyak 73 citra jaringan tumor, 42 citra jaringan stroma, 41 citra jaringan yang tumor yang kompleks, 53 citra jaringan *lympho*, 57 citra jaringan debris, 59 citra jaringan mucosa, 61 citra jaringan *adipose* dan 62 citra jaringan *empty* atau bukan tumor. Terdapat beberapa kesalahan deteksi citra dimana seharusnya citra yang terdeteksi adalah jenis jaringan tumor akan tetapi ditunjukkan sebagai jenis jaringan *lympho*. Berdasarkan table 4.5 *confusion matrix* dapat dihitung nilai akurasi, presisi, *recall (sensitivity)*, secara manual sebagai berikut:

$$\begin{aligned}
 \text{Akurasi} &= \frac{\sum TP}{\text{Total Data}} \times 100\% \\
 &= \frac{TP(0) + TP(1) + TP(2) + TP(3) + TP(4) + TP(5) + TP(6) + TP(7)}{\text{Total Data}} \times 100\% \\
 &= \frac{73 + 42 + 41 + 53 + 57 + 59 + 61 + 62}{500} \times 100\% \\
 &= \frac{448}{500} \times 100\% \\
 &= 89,60\%
 \end{aligned}$$

$$\begin{aligned}
 \text{All Recall} &= \sum \frac{TP}{(TP + FN)} \\
 &= \frac{TP(0)}{(TP(0)+FN(0))} + \frac{TP(1)}{(TP(1)+FN(1))} + \frac{TP(2)}{(TP(2)+FN(2))} + \frac{TP(3)}{(TP(3)+FN(3))} + \frac{TP(4)}{(TP(4)+FN(4))} + \\
 &\quad \frac{TP(5)}{(TP(5)+FN(5))} + \frac{TP(6)}{(TP(6)+FN(6))} + \frac{TP(7)}{(TP(7)+FN(7))} \\
 &= \frac{73}{(73+2)} + \frac{42}{(42+5)} + \frac{41}{(41+17)} + \frac{53}{(53+2)} + \frac{57}{(57+17)} + \frac{59}{(59+6)} + \frac{61}{(61+3)} + \frac{62}{(62+0)} \\
 &= 0,97 + 0,89 + 0,70 + 0,96 + 0,77 + 0,90 + 0,95 + 1 \\
 &= 7,14
 \end{aligned}$$

$$\begin{aligned}
 \text{Recall} &= \frac{R(0) + R(1) + R(2) + R(3) + R(4) + R(5) + R(6) + R(7)}{\text{Jumlah Kelas}} \times 100\% \\
 &= \frac{0,97 + 0,89 + 0,70 + 0,96 + 0,77 + 0,90 + 0,95 + 1}{8} \times 100\% \\
 &= \frac{7,14}{8} \times 100\% \\
 &= 89,25\%
 \end{aligned}$$

$$\begin{aligned}
\text{All Precision} &= \sum \frac{TP}{(TP + FP)} \\
&= \frac{TP(0)}{(TP(0)+FP(0))} + \frac{TP(1)}{(TP(1)+FP(1))} + \frac{TP(2)}{(TP(2)+FP(2))} + \frac{TP(3)}{(TP(3)+FP(3))} + \frac{TP(4)}{(TP(4)+FP(4))} + \\
&\quad \frac{TP(5)}{(TP(5)+FP(5))} + \frac{TP(6)}{(TP(6)+FP(6))} + \frac{TP(7)}{(TP(7)+FP(7))} \\
&= \frac{73}{(73+9)} + \frac{42}{(42+16)} + \frac{41}{(41+7)} + \frac{53}{(53+9)} + \frac{57}{(57+3)} + \frac{59}{(59+4)} + \frac{61}{(61+2)} + \frac{62}{(62+2)} \\
&= 0,89 + 0,72 + 0,85 + 0,85 + 0,95 + 0,93 + 0,96 + 0,96 \\
&= 7,11 \\
\text{Precision} &= \frac{P(0) + P(1) + P(2) + P(3) + P(4) + P(5) + P(6) + P(7)}{\text{Jumlah Kelas}} \times 100\% \\
&= \frac{0,89 + 0,72 + 0,85 + 0,85 + 0,95 + 0,93 + 0,96 + 0,96}{8} \times 100\% \\
&= \frac{7,14}{8} \times 100\% \\
&= 88,87\% \\
\text{All F1-Score} &= \sum \frac{2TP}{(2TP + FP + FN)} \\
&= \frac{2TP(0)}{(2TP(0)+FP(0)+FN(0))} + \frac{2TP(1)}{(2TP(1)+FP(1)+FN(1))} + \frac{2TP(2)}{(2TP(2)+FP(2)+FN(2))} + \\
&\quad \frac{2TP(3)}{(2TP(3)+FP(3)+FN(3))} + \frac{2TP(4)}{(2TP(4)+FP(4)+FN(4))} + \frac{2TP(5)}{(2TP(5)+FP(5)+FN(5))} + \\
&\quad \frac{2TP(6)}{(2TP(6)+FP(6)+FN(6))} + \frac{2TP(7)}{(2TP(7)+FP(7)+FN(7))} \\
&= \frac{2(73)}{(2(73)+9+2)} + \frac{2(42)}{(2(42)+16+5)} + \frac{2(41)}{(2(41)+7+17)} + \frac{2(53)}{(2(53)+9+2)} + \frac{2(57)}{(2(57)+3+17)} + \\
&\quad \frac{2(59)}{(2(59)+4+6)} + \frac{2(61)}{(2(61)+2+3)} + \frac{2(62)}{(2(62)+2+0)} \\
&= \frac{146}{156} + \frac{84}{105} + \frac{82}{106} + \frac{106}{117} + \frac{114}{134} + \frac{118}{128} + \frac{122}{127} + \frac{124}{126} \\
&= 0,93 + 0,80 + 0,77 + 0,90 + 0,85 + 0,92 + 0,96 + 0,98 \\
&= 7,11 \\
\text{F1-Score} &= \frac{F(0) + F(1) + F(2) + F(3) + F(4) + F(5) + F(6) + F(7)}{\text{Jumlah Kelas}} \times 100\% \\
&= \frac{0,93 + 0,80 + 0,77 + 0,90 + 0,85 + 0,92 + 0,96 + 0,98}{8} \times 100\% \\
&= \frac{7,14}{8} \times 100\% \\
&= 88,87\%
\end{aligned}$$

Berikut merupakan tabel rangkuman nilai *recall*, *precision* dan *f1-score* pada setiap kelas dataset kanker kolorektal.

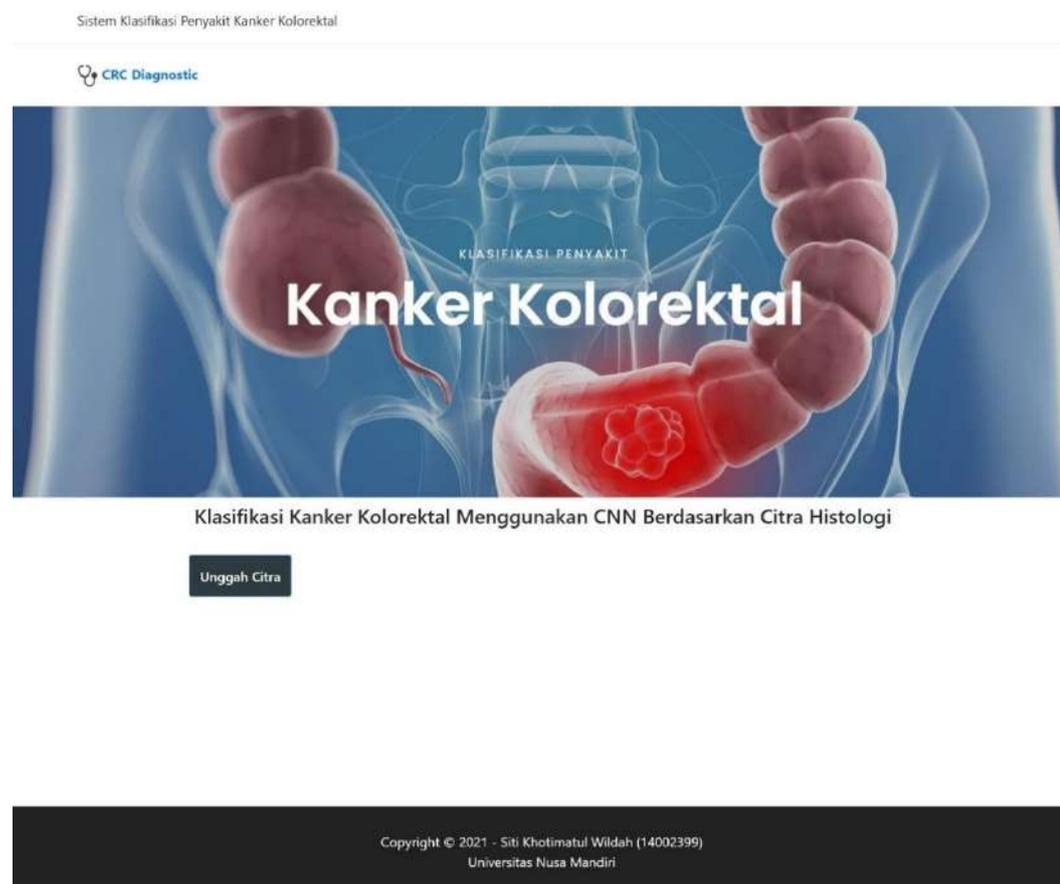
Tabel 4.12. Laporan Klasifikasi

Kelas	Recall	Precision	F1-Score
Tumor (0)	0,97	0,89	0,93
Stroma (1)	0,89	0,72	0,80

Complex (2)	0,70	0,85	0,77
Lympho (3)	0,96	0,85	0,90
Debris (4)	0,77	0,95	0,85
Mucosa (5)	0,90	0,93	0,92
Adipose (6)	0,95	0,96	0,96
Empty (7)	1.00	0,96	0,98

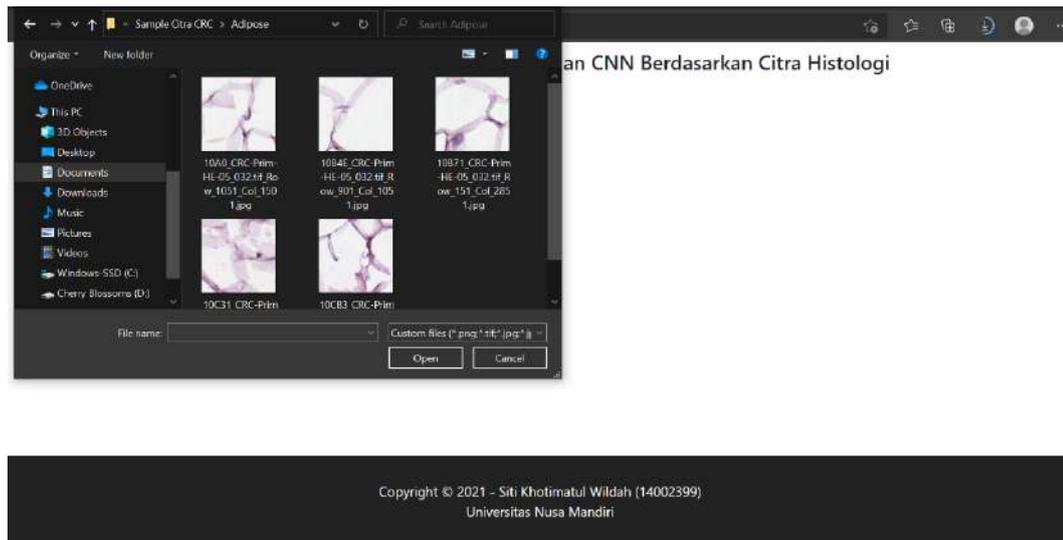
4.3. Implementasi Sistem

Dari hasil pengujian yang dilakukan pada penelitian dengan arsitektur CNN yang diusulkan, diperoleh model algoritma yang dapat digunakan untuk diimplementasikan kedalam sebuah sistem program. Program yang dibuat menggunakan Bahasa pemrograman *python* dengan *framework* yang digunakan adalah *Flask Python Web Development*. Berikut merupakan tampilan antar muka program yang dapat dilihat pada gambar 4.5.



Gambar 4.5 Tampilan awal program

Gambar 4.5 merupakan tampilan awal program, user pada tahap ini hanya perlu melakukan upload citra jaringan kanker kolorektal yang akan diprediksi, seperti pada gambar 4.6.



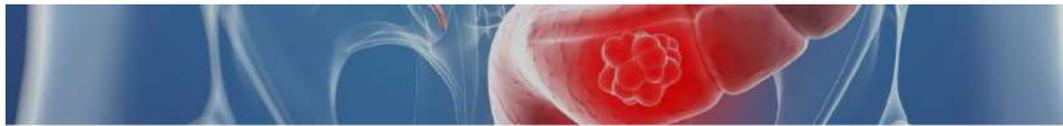
Gambar 4.6 Pemilihan Citra

Setelah melakukan proses upload citra user hanya perlu menekan tombol *predict* pada sistem, maka sistem akan memproses citra masukan tersebut.



Gambar 4.7 Proses Input Citra

Gambar 4.7 merupakan contoh pengujian terhadap sistem yang dibuat untuk mengklasifikasikan jenis jaringan yang ada pada kanker kolorektal berdasarkan citra histologi, dalam hal ini pengujian dilakukan dengan menggunakan salah satu citra pada kelas Adipose, kemudian klik *predict* untuk melihat hasil prediksi yang dihasilkan dari sistem yang telah dibuat.



Klasifikasi Kanker Kolorektal Menggunakan CNN Berdasarkan Citra Histologi

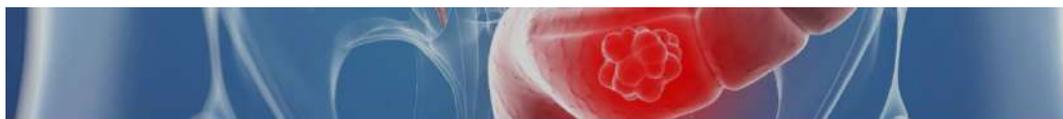
Unggah Citra



Result: Dinyatakan sebagai adiposa, dimana gejala ini disebabkan karena jaringan adiposa menghasilkan banyak hormon dan sitokin pro-inflamasi diantaranya yaitu, interleukin 6, Tumor Necrotic Factor α (TNF- α), leptin dan adiponektin yang memberikan lingkungan mikro inflamasi yang dibutuhkan sebagai tempat hidup sel kanker.

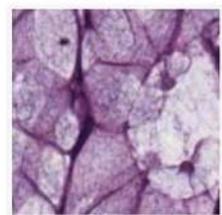
Gambar 4.8 Hasil Prediksi Kelas Jaringan Adipose

Gambar 4.8 menunjukkan bahwa hasil prediksi yang dilakukan terhadap salah satu citra pada kelas jaringan Adipose menghasilkan prediksi yang sesuai dengan jenis citra yang digunakan.



Klasifikasi Kanker Kolorektal Menggunakan CNN Berdasarkan Citra Histologi

Unggah Citra



Result: Dinyatakan sebagai Mucosa yang merupakan sel-sel kanker masih sebatas pada lapisan dinding usus besar (lapisan mukosa). Jenis kanker ini termasuk kategori kanker kolorektal stadium 1.

Gambar 4.9 Hasil Prediksi Kelas Jaringan Mucosa

Gambar 4.9 menunjukkan hasil yang diperoleh dari pengujian pada sistem. Berdasarkan pengujian tersebut, dapat dilihat bahwa prediksi yang dihasilkan oleh sistem sudah sesuai dengan jenis citra dari jaringan kelas yang dijadikan sebagai contoh kasus.

BAB V

KESIMPULAN

Bab ini merupakan bab terakhir pada penelitian yang dilakukan dan berisi mengenai kesimpulan dari hasil yang didapat pada penelitian serta saran-saran yang diusulkan bagi peneliti selanjutnya sebagai pertimbangan pengembangan penelitian.

5.1. Kesimpulan

Berdasarkan hasil penelitian yang dilakukan mengenai Klasifikasi Penyakit Kanker Kolorektal Menggunakan Metode *Convolutional Neural Network* Berdasarkan Citra Histologi, maka dapat ditarik beberapa kesimpulan pada penelitian ini, yaitu sebagai berikut:

1. Pemodelan dengan metode CNN menggunakan arsitektur yang dimodifikasi yakni penggunaan 4 layer konvolusi, ukuran filter 16 sampai dengan 128, dan kernel ukuran 2x2 serta 2 layer *flatten* memberikan kinerja yang baik dengan nilai akurasi yang dihasilkan cukup tinggi dibandingkan dengan penggunaan 1 layer, 2 layer, 3 layer ataupun 5 layer konvolusi.
2. Penerapan hyperparameter pada pemodelan CNN yang diusulkan dapat memberikan hasil yang baik dimana pada penelitian ini digunakan *optimizer* RMSProp, jenis loss *categorical_crossentropy* memberikan hasil yang baik dibandingkan penggunaan *optimizer Adam, SGD, Adadelta* maupun *Adagrad*.
3. Penggunaan persentase jumlah dataset yang digunakan untuk data latih dan data uji dengan rasio 90:10 memberikan hasil yang lebih baik dibandingkan rasio 60:40, rasio 70:30 maupun 80:20.
4. Hasil percobaan penerapan *epoch* 100 dengan *batch size* 50 serta ukuran resolusi citra sebesar 50x50 menunjukkan kinerja yang baik terbukti dengan nilai akurasi yang ditampilkan mengalami kenaikan.
5. Metode CNN yang diusulkan dapat digunakan untuk klasifikasi jenis jaringan penyakit kanker kolorektal dan implementasi model CNN yang diusulkan menjadi sebuah sistem berbasis web dapat mengklasifikasikan jenis jaringan penyakit kanker kolorektal tersebut dengan benar.

5.2. Saran-Saran

Berdasarkan hasil yang diperoleh pada penelitian ini, terdapat beberapa saran yang dapat dijadikan sebagai pertimbangan untuk pengembangan penelitian dimasa yang akan datang:

1. Memperbanyak data atau melakukan augmentasi citra sehingga sistem dapat mempelajari citra lebih banyak dan model yang diusulkan dapat lebih baik.
2. Melakukan penambahan langkah pra pemrosesan citra agar citra yang digunakan lebih baik dan lebih jelas.
3. Memperbaharui arsitektur CNN serta percobaan *hyperparameter* lain yang dimaksudkan untuk memperoleh nilai akurasi yang lebih baik.
4. Menambahkan algoritma lain dalam identifikasi sehingga dapat dilakukan perbandingan hasil identifikasi yang lebih banyak.
5. Melakukan penambahan algoritma lain sehingga dapat dilakukan perbandingan hasil klasifikasi jenis jaringan pada kanker kolorektal yang lebih banyak.

DAFTAR PUSTAKA

- [1] A. Wasil, S. Bakri, Y. Wisnu, and A. Mughni, "Comparison Between Laparotomy And Laparoscopy Colorectal Cancer For Length Of Hospital, Surgical Site Infection and Intraperitoneum Adhesion," vol. 9, no. November, pp. 483–488, 2020.
- [2] L. Zorron Cheng Tao Pu *et al.*, "Computer-aided diagnosis for characterization of colorectal lesions: comprehensive software that includes differentiation of serrated lesions," *Gastrointest. Endosc.*, vol. 92, no. 4, pp. 891–899, 2020, doi: 10.1016/j.gie.2020.02.042.
- [3] P. J. C. Devi, N. P. Sriwidayani, and N. P. Ekawati, "Karakteristik Klinikopatologi Pasien Tumor Kolorektal Yang Menjalani Biopsi Kolonoskopi Di Rsup Sanglah Denpasar Tahun 2015-2017," *J. Kedokt. Indones.*, vol. 9, no. 2, pp. 56–60, 2020, [Online]. Available: <https://www.jurnalmedika.com/blog/124-Retensio-Urine-Post-Partum>.
- [4] U. Athiyah, I. Muhimmah, and E. Marfianti, "Ekstraksi Ciri Polip dan Pendarahan Berdasarkan Citra Endoskopi Kolorektal," *J. Inform. J. Pengemb. IT*, vol. 3, no. 1, pp. 81–85, 2018.
- [5] F. Felix, J. Wijaya, S. P. Sutra, P. W. Kosasih, and P. Sirait, "Implementasi Convolutional Neural Network Untuk Identifikasi Jenis Tanaman Melalui Daun," *J. SIFO Mikroskil*, vol. 21, no. 1, pp. 1–10, 2020.
- [6] I. Bagus, L. Mahadya, R. S. Hartati, and Y. Divayana, "Diagnosa Tumor Otak Berdasarkan Citra MRI (Magnetic Resonance Imaging)," vol. 18, no. 2, 2019.
- [7] N. Sakinah *et al.*, "Analisis Kinerja Algoritma Mesin Pembelajaran Untuk Klasifikasi Penyakit Stroke Menggunakan Citra Ct Scan," vol. 7, no. 4, pp. 833–844, 2020, doi: 10.25126/jtiik.202073482.
- [8] Muhammad Luthfi Adnan, "Potensi Kombinasi Bakteri Probiotik *Lactobacillus reuteri* dengan Sari Buah Kurma (*Phoenix dactylifera*) sebagai Terapi Penyakit Kanker Kolorektal," *Scr. SCORE Sci. Med. J.*, vol. 1, no. 2, p. 11, 2020, doi: 10.32734/scripta.v1i2.1233.
- [9] J. Jumadi, A. Al Akbar, and S. E. Septiawan, "Implementasi Metode Template Matching Untuk Klasifikasi Citra Anggrek Pensil Bengkulu," no. June, 2021.
- [10] A. Gaghaube, M. M. Kaseke, and S. J. R. Kalangi, "Karakteristik Gambaran Histologis Paru-Paru Pasien COVID-19," *eBiomedik*, vol. 9, no. 1, pp. 52–67, 2021, doi: 10.35790/ebm.9.1.2021.31896.
- [11] P. A. Nugroho, I. Fenriana, R. Ariyanto, and M. Kom, "Implementasi Deep Learning Menggunakan Convolutional Neural Network (Cnn) Pada Ekspresi Manusia," *Implementasi Deep Learn. Menggunakan Convolutional Neural Netw. (Cnn) Pada Ekspresi Mns.*, vol. 2, no. 1, 2020.
- [12] Ekojono, F. Rahutomo, and D. N. Sari, "Implementasi Library Deep Learning Keras pada Sistem Ujian Essay Online," *J. Inform. Polinema*, vol. 6, no. 2, pp. 73–79, 2020, doi: 10.33795/jip.v6i2.303.
- [13] E. Maggiori, Y. Tarabalka, G. Charpiat, P. Alliez, and S. Member, "Convolutional Neural Networks for Large-Scale Remote Sensing Image Classification IEEE Transactions On Geoscience And Remote Sensing 1

- Convolutional Neural Networks for Large-Scale Remote Sensing Image Classification,” *Ieee Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 645–657, 2017, [Online]. Available: <https://hal.inria.fr/hal-01369906>.
- [14] M. Afif, A. Fawwaz, K. N. Ramadhani, and F. Sthevanie, “Klasifikasi Ras pada Kucing menggunakan Algoritma Convolutional Neural Network (CNN),” *e-Proceeding Eng.*, vol. 8, no. 1, pp. 715–730, 2021.
- [15] P. N. Candra and A. Prapanca, “Klasifikasi gambar asli dan manipulasi menggunakan Error Level Analysis (ELA) sebagai proses komputasi metode Convolutional Neural Network (CNN),” *J. Informatics Comput. Sci.*, vol. 02, pp. 9–18, 2020.
- [16] D. D. Saputra, B. Pratama, Y. Akbar, W. Gata, P. Studi, and M. Ilmu, “Penerapan text mining untuk assingment complaint handling customer terhadap divisi terkait menggunakan metode decision tree algoritma c4.5 (studi case : pt. xl axiata, tbk),” vol. 11, no. 2, pp. 207–216, 2018.
- [17] R. Siringoringo, “Klasifikasi Data Tidak Seimbang Menggunakan Algoritma SMOTE dan k-Nearest Neighbor,” *J. ISD*, vol. 3, no. 1, pp. 44–49, 2018.
- [18] R. Yotenka and F. El Huda, “View of Implementasi Long Short-Term Memory Pada Harga Saham Perusahaan Perkebunan Di Indonesia.pdf.” 2020.
- [19] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–15, 2015.
- [20] A. C. Sitepu and M. Sigiyo, “Analisis Fungsi Aktivasi Relu Dan Sigmoid Menggunakan Optimizer Sgd Dengan Representasi Mse Pada Model Backpropagation,” pp. 12–25, 2021.
- [21] A. A. Lydia and F. S. Francis, “Adagrad-An Optimizer for Stochastic Gradient Descent,” *Int. J. Inf. Comput. Sci.*, vol. 6, no. 5, pp. 566–568, 2019, [Online]. Available: <http://ijics.com>.
- [22] S. Ruder, “An overview of gradient descent optimization algorithms,” pp. 1–14, 2017, [Online]. Available: <http://arxiv.org/abs/1609.04747>.
- [23] JUD and J. Enterprise, *Mastering Python*. 2017.
- [24] G. Subramanian, *Python Data Science Cookbook*. Birmingham: Packt Publishing Ltd., 2015.
- [25] R. Irsyad, “Penggunaan Python Web Framework Flask Untuk Pemula,” 2018, doi: 10.31219/osf.io/t7u5r.
- [26] S. F. Pane and E. V. Rahchmadani, *Big Data: Forecasting Menggunakan Python*. Bandung: Kreatif Industri Nusantara, 2020.
- [27] L. N. Rizalputri, T. Pranata, N. S. Tanjung, H. M. Auliya, S. Harimurti, and I. Anshori, “Colorectal Histology CSV Multi-classification Accuracy Comparison using Various Machine Learning Models,” *Proc. Int. Conf. Electr. Eng. Informatics*, vol. 2019-July, no. October 2020, pp. 58–62, 2019, doi: 10.1109/ICEEI47359.2019.8988846.
- [28] K. Zhang *et al.*, “Multiple Morphological Constraints-Based Complex Gland Segmentation in Colorectal Cancer Pathology Image Analysis,” *Complexity*, vol. 2020, 2020, doi: 10.1155/2020/6180457.
- [29] N. Sengar, N. Mishra, M. K. Dutta, J. Prinosil, and R. Burget, “Grading of colorectal cancer using histology images,” *2016 39th Int. Conf. Telecommun. Signal Process. TSP 2016*, pp. 529–532, 2016, doi:

- 10.1109/TSP.2016.7760936.
- [30] D. Sarwinda, R. H. Paradisa, A. Bustamam, and P. Anggia, "Deep Learning in Image Classification using Residual Network (ResNet) Variants for Detection of Colorectal Cancer," *Procedia Comput. Sci.*, vol. 179, no. 2019, pp. 423–431, 2021, doi: 10.1016/j.procs.2021.01.025.
- [31] J. N. Kather *et al.*, "Multi-class texture analysis in colorectal cancer histology," *Sci. Rep.*, vol. 6, pp. 1–11, 2016, doi: 10.1038/srep27988.
- [32] G. De Pietro, L. Gallo, R. J. Howlett, and L. C. Jain, "Dimensionality Reduction Strategies for CNN-Based Classification of Histopathological Images Silvia," *Intell. Interact. Multimed. Syst. Serv. 2017*, vol. 1, 2018, doi: 10.1007/978-3-319-59480-4.
- [33] M. Shapcott, K. J. Hewitt, and N. Rajpoot, "Deep learning with sampling in colon cancer histology," *Front. Bioeng. Biotechnol.*, vol. 7, no. MAR, 2019, doi: 10.3389/fbioe.2019.00052.
- [34] M. Morkunas, P. Treigys, J. Bernatavičienė, A. Laurinavičius, and G. Korvel, "Machine learning based classification of colorectal cancer tumour tissue in whole-slide images," *Inform.*, vol. 29, no. 1, pp. 75–90, 2018, doi: 10.15388/Informatica.2018.158.
- [35] B. Zhang, X. L. Liang, H. Y. Gao, L. S. Ye, and Y. G. Wang, "Models of logistic regression analysis, support vector machine, and back-propagation neural network based on serum tumor markers in colorectal cancer diagnosis," *Genet. Mol. Res.*, vol. 15, no. 2, 2016, doi: 10.4238/gmr.15028643.
- [36] Y. Pan *et al.*, "Screening and diagnosis of colorectal cancer and advanced adenoma by Bionic Glycome method and machine learning," vol. 11, no. 6, pp. 3002–3020, 2021.
- [37] J. Gründner, H. U. Prokosch, M. Stürzl, R. Croner, J. Christoph, and D. Toddenroth, "Predicting clinical outcomes in colorectal cancer using machine learning," *Stud. Health Technol. Inform.*, vol. 247, pp. 101–105, 2018, doi: 10.3233/978-1-61499-852-5-101.
- [38] X. Yue, N. Dimitriou, P. D. Caie, D. J. Harrison, and O. Arandjelović, "Colorectal cancer outcome prediction from H&E whole slide images using machine learning and automatically inferred phenotype profiles," *Proc. 11th Int. Conf. Bioinforma. Comput. Biol. BiCOB 2019*, vol. 60, pp. 139–149, 2019, doi: 10.29007/n912.
- [39] N. Kaoungku, K. Kerdprasop, N. Kerdprasop, R. Chanklan, and K. Suksut, "Colorectal Cancer Histology Image Classification Using Stacked Ensembles," *Int. J. Futur. Comput. Commun.*, vol. 8, no. 3, pp. 104–108, 2019, doi: 10.18178/ijfcc.2019.8.3.549.
- [40] O. J. Skrede *et al.*, "Deep learning for prediction of colorectal cancer outcome: a discovery and validation study," *Lancet*, vol. 395, no. 10221, pp. 350–360, 2020, doi: 10.1016/S0140-6736(19)32998-8.
- [41] N. Dimitriou, O. Arandjelović, D. J. Harrison, and P. D. Caie, "A principled machine learning framework improves accuracy of stage II colorectal cancer prognosis," *npj Digit. Med.*, vol. 1, no. 1, pp. 1–9, 2018, doi: 10.1038/s41746-018-0057-x.

DAFTAR RIWAYAT HIDUP

A. Biodata Mahasiswa

Nama Lengkap : Siti Khotimatul Wildah
NIM : 14002399
Tempat & Tanggal Lahir : Sukabumi, 10 September 1997
Jenis Kelamin : Perempuan
Alamat : Kp. Lio RT 002 RW 001 Desa Cireunghas
Kecamatan Cireunghas, Kabupaten Sukabumi,
Provinsi Jawa Barat, 43193
Nomor Telepon : 0896-5770-7497

Riwayat Pendidikan Formal & Non-Formal

1. SDN Lio, lulus tahun 2009
2. SMPN 1 Sukaraja, lulus tahun 2012
3. SMK Pasim Plus Sukabumi, lulus tahun 2015
4. Universitas Bina Sarana Informatika, lulus tahun 2019

C. Riwayat Pengalaman Berorganisasi / Pekerjaan

1. Anggota Pramuka SDN Lio 2007/2009.
2. Anggota PMR SMPN 1 Sukaraja tahun Akademik 2009/2012.
3. Ketua Sekretaris Bidang 5 Osis SMPN 1 Sukaraja tahun Akademik 2011/2012.
4. Anggota Paskibra SMK Pasim Plus Sukabumi tahun Akademik 2012/2015.
5. Praktek Kerja Lapangan di CV SALAFINDO COMPUTER Sukabumi tahun 2014
6. Sekretaris HIMMI AMIK BSI Sukabumi periode 2016-2017
7. PKL/ Riset di PT. Pos Indonesia Sukabumi 2017
8. Asisten Instruktur di Universitas Bina Sarana Informatika 2017
9. Magang sebagai Pustakawati 2018 – 2019
10. Staf Warek II Bidang Non Akademik dan Pengajar di Universitas Bina Sarana Informatika (2019 - sekarang)



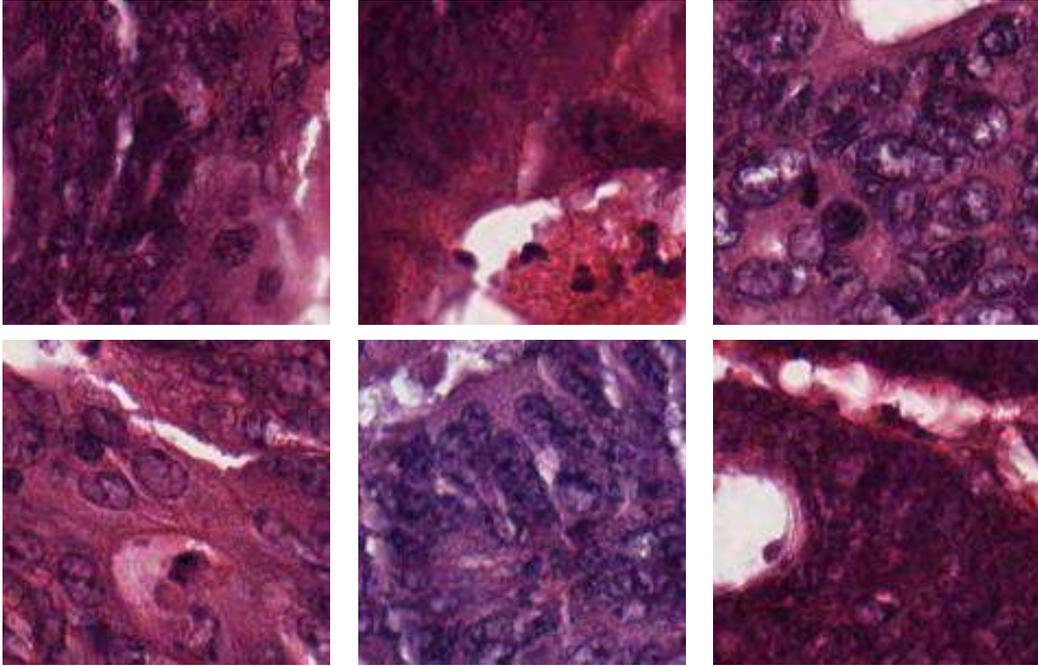
Jakarta, 2 Agustus 2021

Siti Khotimatul Wildah

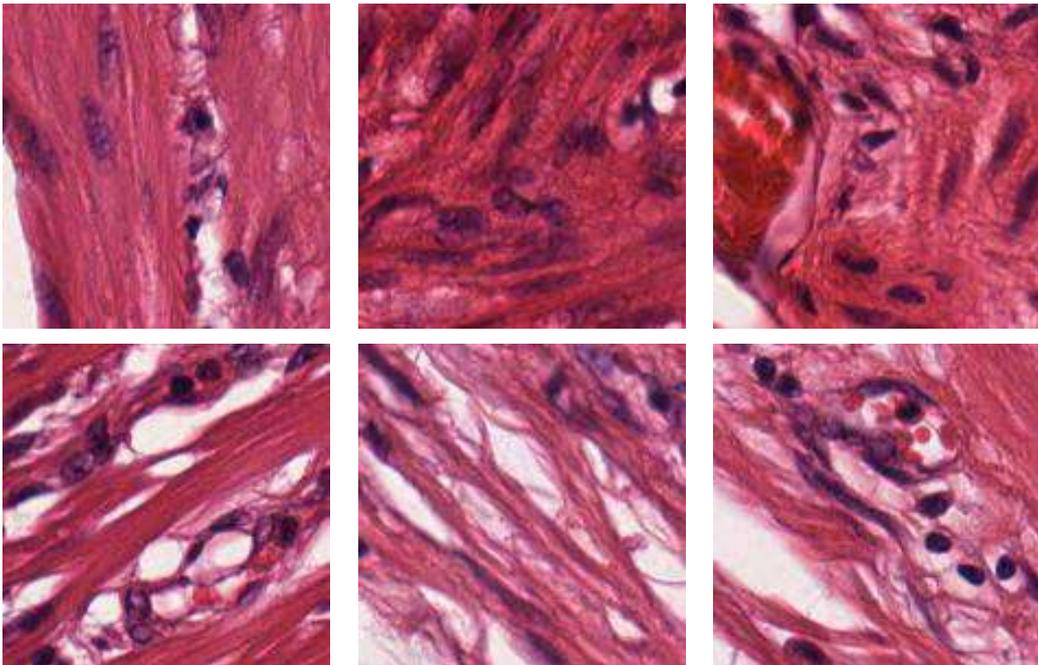
LAMPIRAN

Lampiran 1 Dataset Citra Histologi Kanker Kolorektal

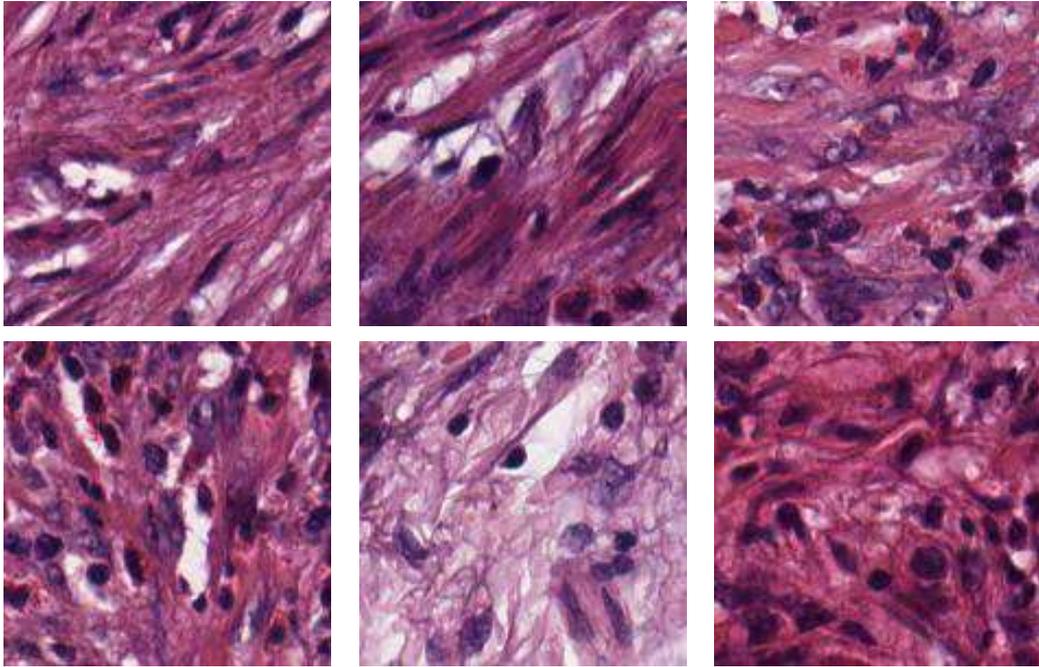
A. Sample Citra Jaringan Tumor



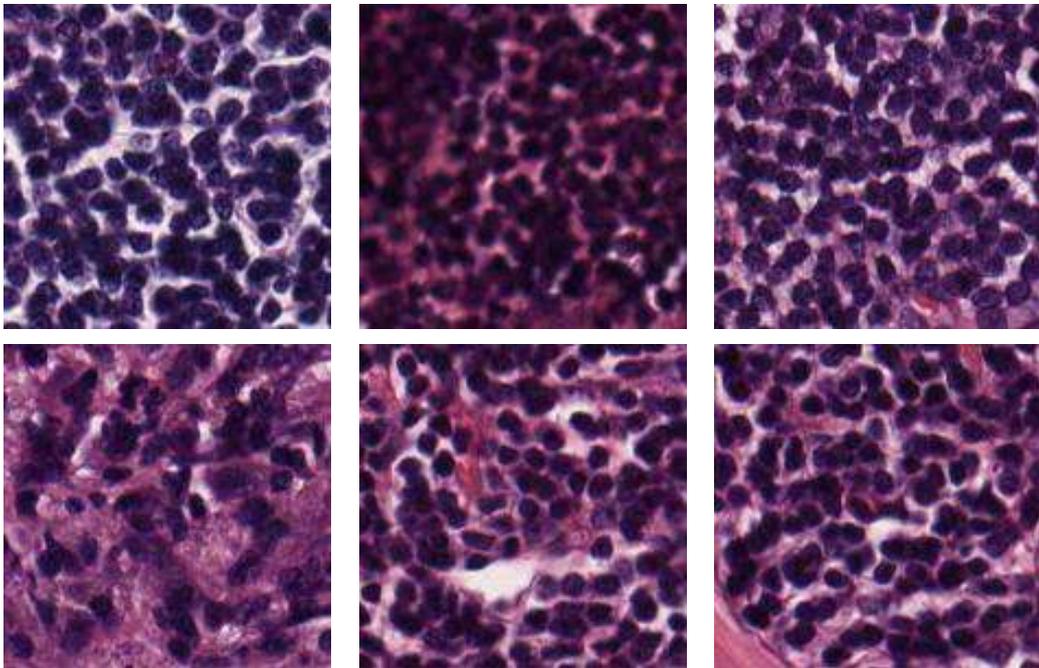
B. Sample Citra Jaringan Stroma



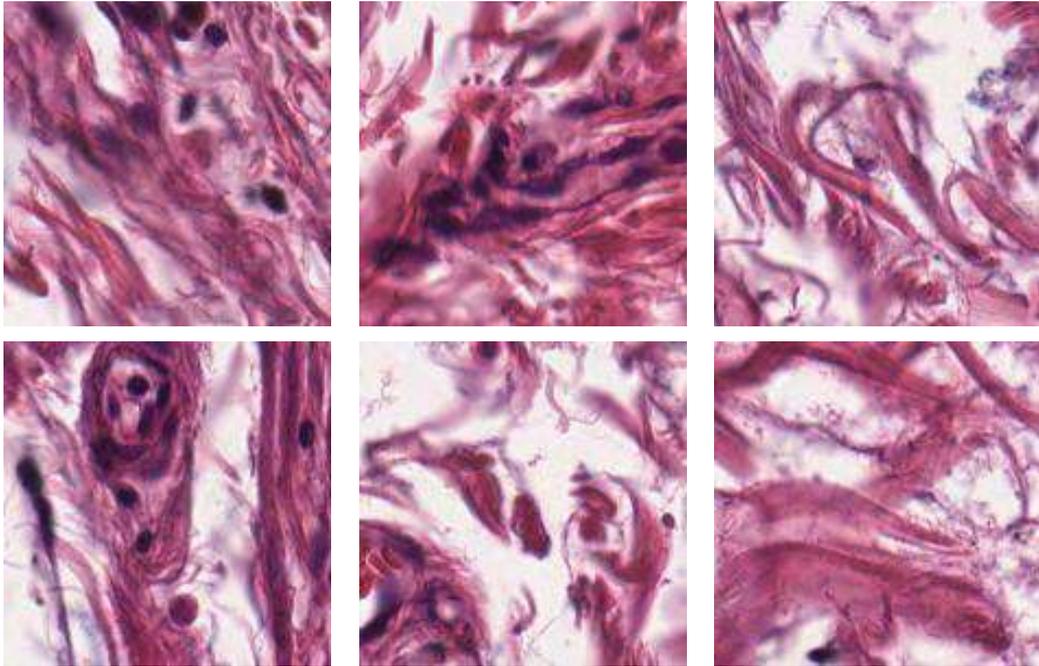
C. Sample Citra Jaringan Complex



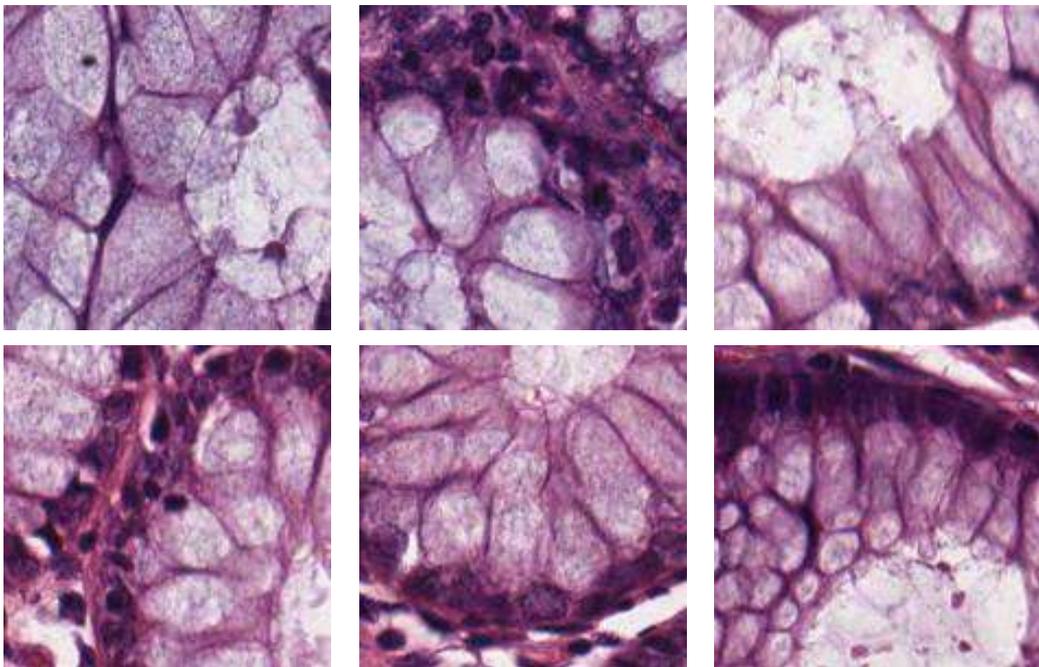
D. Sample Citra Jaringan Lympho



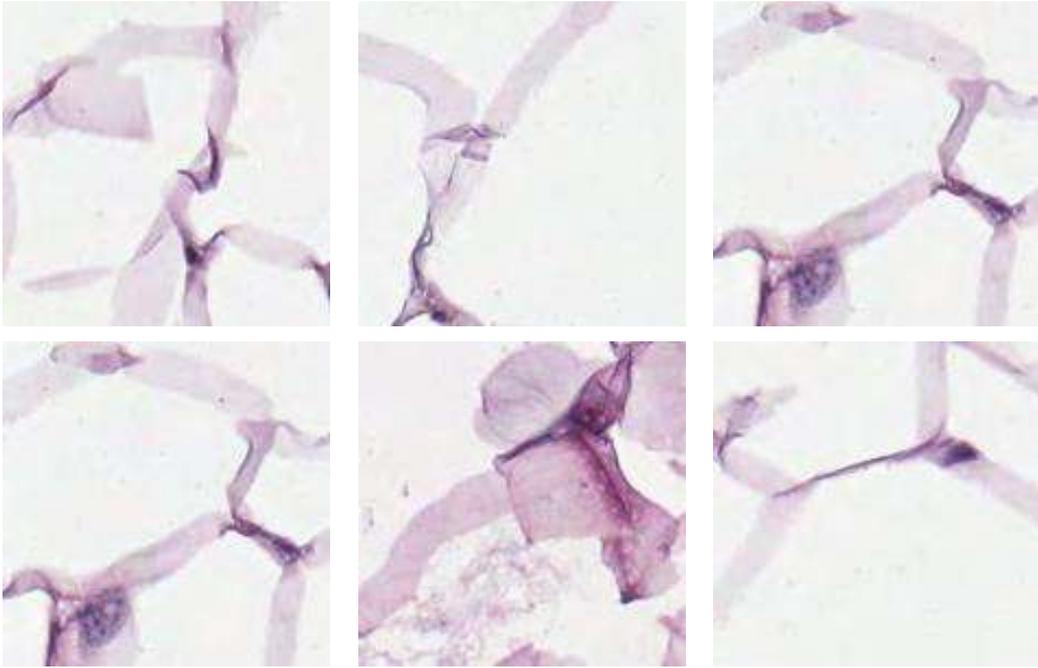
E. Sample Citra Jaringan Debris



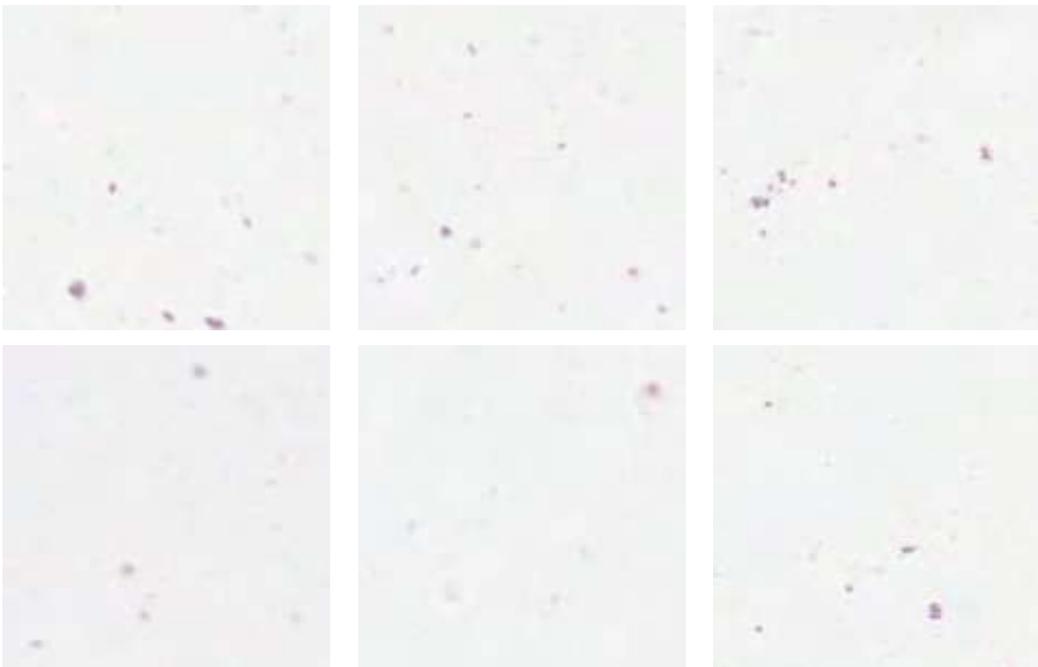
F. Sample Citra Jaringan Mucosa



G. Sample Citra Jaringan Adipose

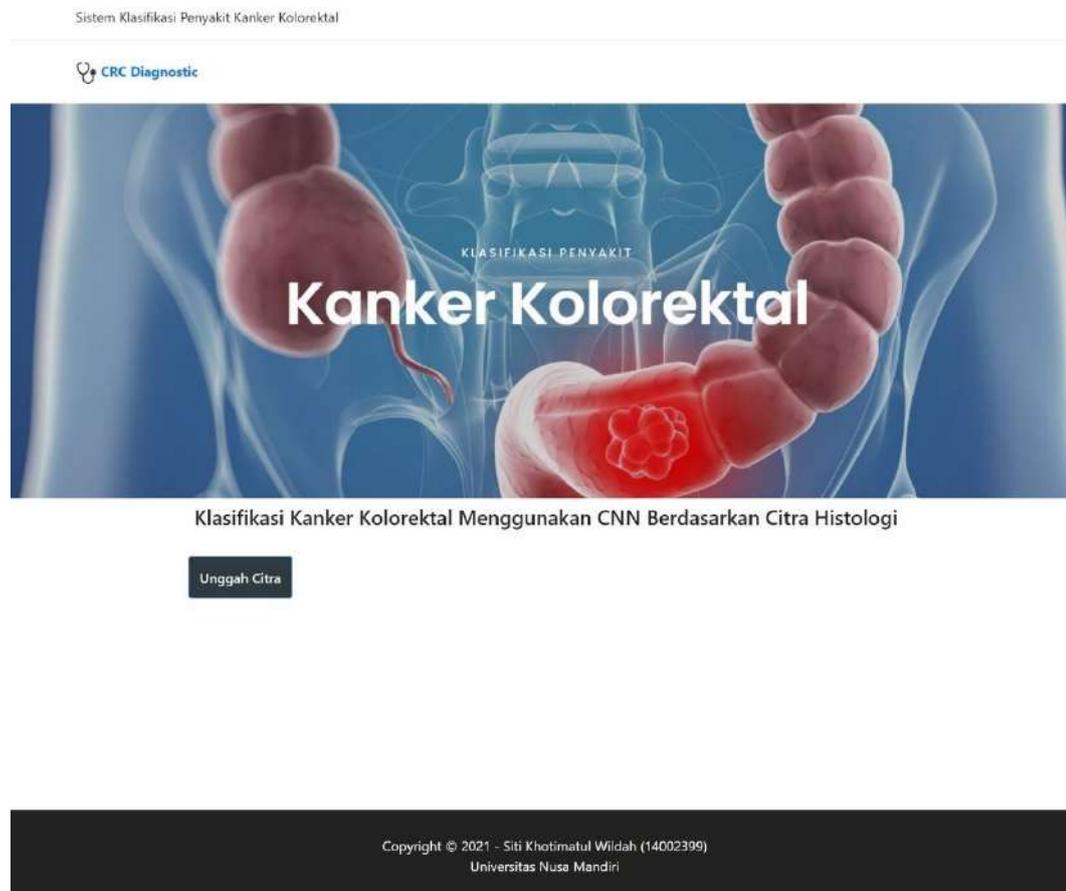


H. Sample Citra Jaringan Adipose

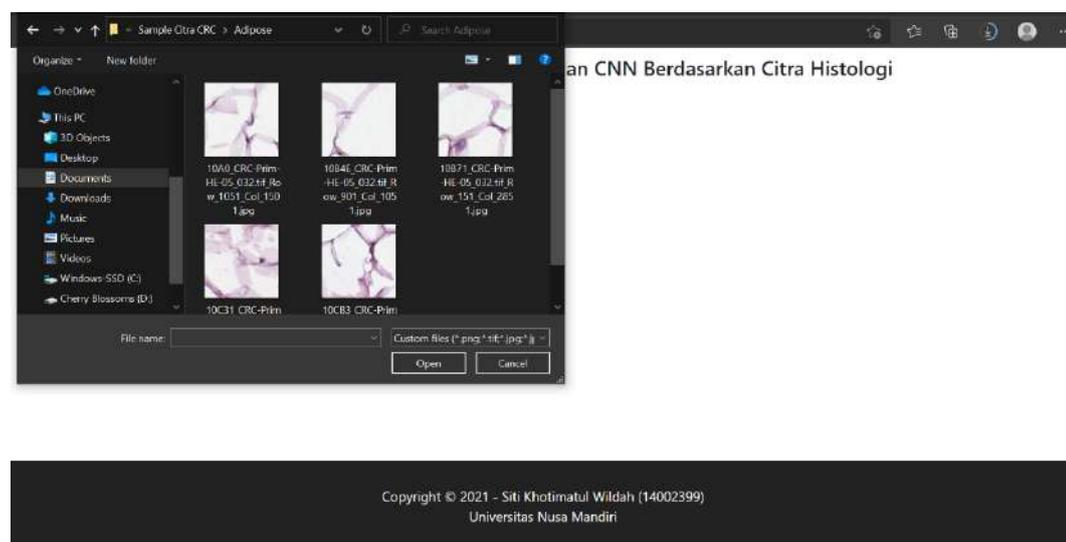


Lampiran 2. Tampilan Antarmuka Program

A. Tampilan Beranda Program



B. Tampilan Pemilihan Citra Histologi



C. Tampilan Citra yang Dipilih



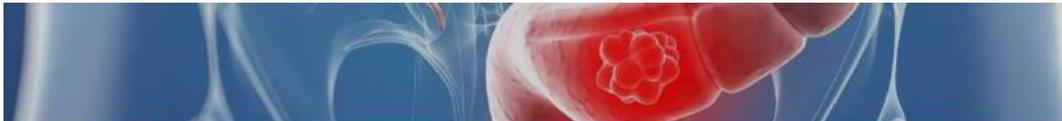
Klasifikasi Kanker Kolorektal Menggunakan CNN Berdasarkan Citra Histologi

Unggah Citra



Prediksi

D. Hasil Prediksi Citra Jaringan Adipose



Klasifikasi Kanker Kolorektal Menggunakan CNN Berdasarkan Citra Histologi

Unggah Citra



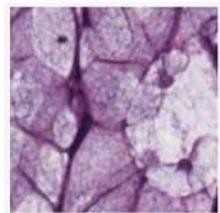
Result: Dinyatakan sebagai adiposa, dimana gejala ini disebabkan karena jaringan adiposa menghasilkan banyak hormon dan sitokin pro-inflamasi diantaranya yaitu, interleukin 6, Tumor Necrotic Factor α (TNF- α), leptin dan adiponektin yang memberikan lingkungan mikro inflamasi yang dibutuhkan sebagai tempat hidup sel kanker.

E. Hasil Prediksi Citra Jaringan Mukosa



Klasifikasi Kanker Kolorektal Menggunakan CNN Berdasarkan Citra Histologi

Unggah Citra



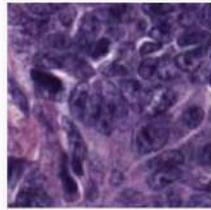
Result: Dinyatakan sebagai Mukosa yang merupakan sel-sel kanker masih sebatas pada lapisan dinding usus besar (lapisan mukosa). Jenis kanker ini termasuk kategori kanker kolorektal stadium 1.

F. Hasil Prediksi Citra Jaringan Tumor



Klasifikasi Kanker Kolorektal Menggunakan CNN Berdasarkan Citra Histologi

Unggah Citra



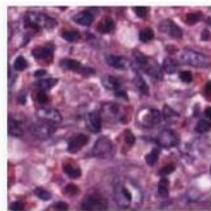
Result: Dinyatakan sebagai tumor maligna yang muncul dari jaringan epitel kolon atau rektum.

G. Hasil Prediksi Citra Jaringan Complex



Klasifikasi Kanker Kolorektal Menggunakan CNN Berdasarkan Citra Histologi

Unggah Citra



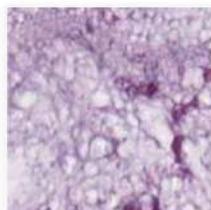
Result: Dinyatakan sebagai Complex atau stroma kompleks yang merupakan jenis kanker kolorektal yang menyerang dinding kolon akan tetapi jenis kanker ini sudah memasuki tingkat keparahan yang cukup tinggi, stroma kompleks mengandung sel tumor tunggal dan atau sedikit sel imun

H. Hasil Prediksi Citra Jaringan Debris



Klasifikasi Kanker Kolorektal Menggunakan CNN Berdasarkan Citra Histologi

Unggah Citra



Result: Dinyatakan sebagai debris yang merupakan puing-puing sisa sel tumor (termasuk nekrosis, perdarahan dan lendir).

I. Hasil Prediksi Citra Jaringan Debris



Klasifikasi Kanker Kolorektal Menggunakan CNN Berdasarkan Citra Histologi

Unggah Citra



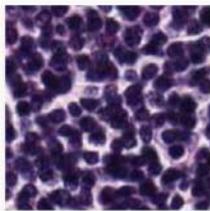
Result: Dinyatakan sebagai Empty atau tidak terdeteksi adanya tumor kanker kolorektal.

J. Hasil Prediksi Citra Jaringan Lympho



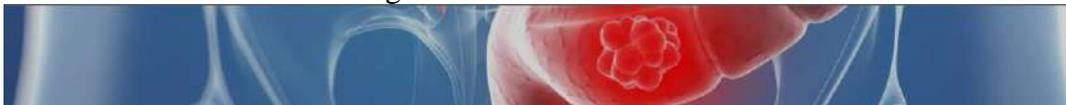
Klasifikasi Kanker Kolorektal Menggunakan CNN Berdasarkan Citra Histologi

Unggah Citra



Result: Dinyatakan sebagai Lympho yang merupakan jenis kanker kolorektal yang telah menembus dinding kolon juga dapat menembus darah atau kelenjar getah bening (lymph vessels).

K. Hasil Prediksi Citra Jaringan Stroma



Klasifikasi Kanker Kolorektal Menggunakan CNN Berdasarkan Citra Histologi

Unggah Citra



Result: Dinyatakan sebagai Stroma, dimana stroma merupakan Jenis kanker yang menyerang sel-sel khusus dinding kolon yang disebut interstitial Cajal.

Lampiran 3. Coding Pengolahan dan Pemodelan CNN

A. Penentuan Direktori Dataset

```
[ ] from google.colab import drive
    drive.mount('/content/drive')
```

Mounted at /content/drive

```
[ ] #Menentukan lokasi path dataset

data_path="/content/drive/My Drive/Colab Notebooks/colorectal/"
file_ext = '.tif'
```

B. Pemanggilan Library Python

```
import os
import pandas as pd
import glob
import cv2
import numpy as np
from matplotlib.colors import Normalize

np.random.seed (42)
%matplotlib inline
import matplotlib.pyplot as plt
import seaborn as sns
plt.rcParams ['figure.figsize'] = (7, 4.5) #Make the figures a bit bigger
import time
import datetime as dt
from sklearn import datasets, metrics
from sklearn.model_selection import train_test_split, cross_val_score
from sklearn.metrics import confusion_matrix, accuracy_score, classification_report
```

C. Merubah Ukuran Citra

```
[ ] # mengambil training labels
labels=os.listdir(data_path)

# mengurutkan training labels
labels.sort()

# membaca dataset citra dan labelnya
ukuranPixImag=tuple((50,50))
data_image = [] # siapkan variabel data_citra berupa sebuah list kosong
data_label = [] # siapkan variabel data_label berupa sebuah list kosong
for nama_klas in labels:
    dir=os.path.join(data_path,nama_klas)
    current_label=nama_klas
    for file in glob.glob(dir + "/*" + file_ext):
        image=cv2.imread(file)
        image=cv2.resize(image,ukuranPixImag)
        data_image.append(image)
        data_label.append(current_label)
data_image = np.array(data_image) # ubah list data_citra ke dalam bentuk numpy array
data_label = np.array(data_label) # ubah list data_label ke dalam bentuk numpy array
```

D. Menentukan Persentase Data Uji dan Data Latih

```
[ ] # Membagi dataset ke dalam data training dan testing
import sklearn.model_selection as model_selection

x_train, x_test, y_train, y_test = model_selection.train_test_split
(data_image, data_label, train_size=0.9, test_size=0.1, random_state=101)

train_data = (x_train, y_train) # Data training (data_citra, data_label)
test_data = (x_test, y_test) # Data testing (data_citra, data_label)

crc = (train_data, test_data) # Dataset kita yang telah terbagi training dan testing
crc = np.array(crc) # Ubah format list dataset_kita menjadi numpy array
```

```
[ ] def pre_process(y):
    y = y.reshape((len(y), 1))
    return y

y_train = pre_process(y_train)
y_test = pre_process(y_test)

print("y_train", y_train.shape)
print("y_test", y_test.shape)

y_train (4500, 1)
y_test (500, 1)
```

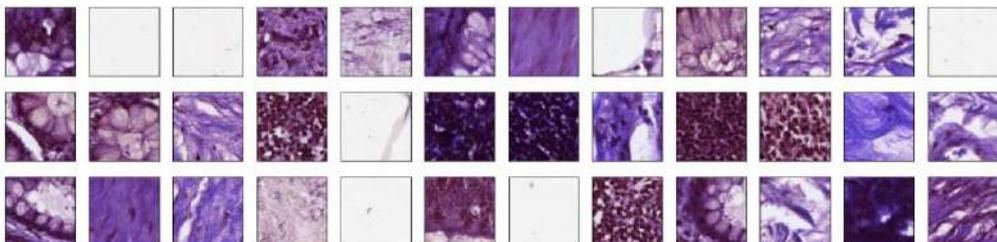
```
[ ] print("X_train original shape", x_train.shape)
print("y_train original shape", y_train.shape)
print("X_test original shape", x_test.shape)
print("y_test original shape", y_test.shape)

X_train original shape (4500, 50, 50, 3)
y_train original shape (4500, 1)
X_test original shape (500, 50, 50, 3)
y_test original shape (500, 1)
```

E. Menampilkan Citra Setelah Proses Resize Menjadi Ukuran 50x50 pixels

```
[ ] import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline

fig = plt.figure(figsize=(20,5))
for i in range(36):
    ax = fig.add_subplot(3, 12, i + 1, xticks=[], yticks=[])
    ax.imshow(np.squeeze(X_train[i]))
```



F. Merubah Skala Nilai Citra

```
[ ] x_train = x_train.astype('float32')/255
    x_test = x_test.astype('float32')/255
```

G. Pembagian Data Uji, Data Latih dan Data Validasi

```
[ ] import keras
    import keras.utils
    from tensorflow import keras
    from keras import utils as np_utils

    # one-hot encode the labels
    num_classes = 8
    y_train = keras.utils.to_categorical(y_train, num_classes)
    y_test = keras.utils.to_categorical(y_test, num_classes)

    # break training set into training and validation sets
    (x_train, x_valid) = x_train[50:], x_train[:50]
    (y_train, y_valid) = y_train[50:], y_train[:50]

    print('x_train shape:', x_train.shape)
    print(x_train.shape[0], 'train samples')
    print(x_test.shape[0], 'test samples')
    print(x_valid.shape[0], 'validation samples')
```

```
x_train shape: (4450, 50, 50, 3)
4450 train samples
500 test samples
50 validation samples
```

H. Pemodelan Model Arsitektur CNN

```
[ ] from keras.models import Sequential
    from keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout

    model = Sequential()
    #Feature Learning
    # 1st conv
    model.add(Conv2D(filters=16, kernel_size=2, padding='same', activation='relu',
                    input_shape=(50, 50, 3)))
    model.add(MaxPooling2D(pool_size=2))
    # 2nd conv
    model.add(Conv2D(filters=32, kernel_size=2, padding='same', activation='relu'))
    model.add(MaxPooling2D(pool_size=2))
    # 3rd conv
    model.add(Conv2D(filters=64, kernel_size=2, padding='same', activation='relu'))
    model.add(MaxPooling2D(pool_size=2))
    # 4nd conv
    model.add(Conv2D(filters=128, kernel_size=2, padding='same', activation='relu'))
    model.add(MaxPooling2D(pool_size=2))

    model.add(Flatten())
    #fully conected layer 1
    model.add(Dropout(0.2))
    model.add(Dense(200, activation='relu'))
    #fully conected layer 2
    model.add(Dropout(0.3))
    model.add(Dense(500, activation='relu'))

    model.add(Dropout(0.4))
    model.add(Dense(8, activation='softmax'))

    model.summary()
```

I. Arsitektur Model CNN

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 50, 50, 16)	208
max_pooling2d_4 (MaxPooling2D)	(None, 25, 25, 16)	0
conv2d_5 (Conv2D)	(None, 25, 25, 32)	2080
max_pooling2d_5 (MaxPooling2D)	(None, 12, 12, 32)	0
conv2d_6 (Conv2D)	(None, 12, 12, 64)	8256
max_pooling2d_6 (MaxPooling2D)	(None, 6, 6, 64)	0
conv2d_7 (Conv2D)	(None, 6, 6, 128)	32896
max_pooling2d_7 (MaxPooling2D)	(None, 3, 3, 128)	0
flatten_1 (Flatten)	(None, 1152)	0
dropout_3 (Dropout)	(None, 1152)	0
dense_3 (Dense)	(None, 200)	230600
dropout_4 (Dropout)	(None, 200)	0
dense_4 (Dense)	(None, 500)	100500
dropout_5 (Dropout)	(None, 500)	0
dense_5 (Dense)	(None, 8)	4008

Total params: 378,548
Trainable params: 378,548
Non-trainable params: 0

J. Penentuan Hyperparameter

```
[ ] model.compile(loss = 'categorical_crossentropy',  
                  optimizer = 'rmsprop',  
                  metrics = ['accuracy'])
```

K. Pelatihan Model

```
[ ] from keras.callbacks import ModelCheckpoint  
  
# train the model  
checkpointer = ModelCheckpoint(filepath='model_self_4layer_100ep50bs_rmsprop_crc_weights.best.hdf5',  
                               verbose=1, save_best_only=True)  
  
hist = model.fit(x_train, y_train, batch_size=50, epochs=100,  
                validation_data=(x_valid, y_valid), callbacks=[checkpointer],  
                verbose=2, shuffle=True)
```

L. Proses Pelatihan Model

```
Epoch 00073: val_loss did not improve from 0.05610
Epoch 74/100
79/79 - 21s - loss: 0.0602 - accuracy: 0.9210 - val_loss: 0.0925 - val_accuracy: 0.8600

Epoch 00074: val_loss did not improve from 0.05610
Epoch 75/100
79/79 - 21s - loss: 0.0580 - accuracy: 0.9210 - val_loss: 0.1225 - val_accuracy: 0.8200

Epoch 00075: val_loss did not improve from 0.05610
Epoch 76/100
79/79 - 21s - loss: 0.0656 - accuracy: 0.9182 - val_loss: 0.1998 - val_accuracy: 0.8600

Epoch 00076: val_loss did not improve from 0.05610
Epoch 77/100
79/79 - 22s - loss: 0.0559 - accuracy: 0.9225 - val_loss: 0.1100 - val_accuracy: 0.8400

Epoch 00077: val_loss did not improve from 0.05610
Epoch 78/100
79/79 - 22s - loss: 0.0569 - accuracy: 0.9243 - val_loss: 0.1245 - val_accuracy: 0.8000

Epoch 00078: val_loss did not improve from 0.05610
Epoch 79/100
79/79 - 23s - loss: 0.0543 - accuracy: 0.9195 - val_loss: 0.1059 - val_accuracy: 0.8400

Epoch 00079: val_loss did not improve from 0.05610
Epoch 80/100
79/79 - 22s - loss: 0.0506 - accuracy: 0.9306 - val_loss: 0.1145 - val_accuracy: 0.8600
```

M. Evaluasi Model

```
[ ] # evaluate and print test accuracy
score = model.evaluate(x_test, y_test, verbose=0)
print('\n', 'Test accuracy: ', score[1])

[ ] # get predictions on the test set
y_hat = model.predict(x_test)

# define text labels
crc = ['Tumor', 'Stroma', 'Complex', 'Lympho', 'Debris', 'Mucosa', 'Adipose', 'Empty']

[ ] # plot a random sample of test images, their predicted labels, and ground truth
fig = plt.figure(figsize=(20, 8))
for i, idx in enumerate(np.random.choice(x_test.shape[0], size=32, replace=False)):
    ax = fig.add_subplot(4, 8, i + 1, xticks=[], yticks=[])
    ax.imshow(np.squeeze(x_test[idx]))
    pred_idx = np.argmax(y_hat[idx])
    true_idx = np.argmax(y_test[idx])
    ax.set_title("{} ({}).format(crc[pred_idx], crc[true_idx]),
                color=("green" if pred_idx == true_idx else "red"))

[ ] # plot learning curves of model accuracy
plt.plot(hist.history['accuracy'], label='Accuracy')
plt.plot(hist.history['val_accuracy'], label='Val Accuracy')
plt.legend()
plt.show()
```

```
[ ]
plt.plot(hist.history['loss'], label='Loss')
plt.plot(hist.history['val_loss'], label='Val Loss')
plt.legend()
plt.show()
```

```
[ ] # plot learning curves of model accuracy
plt.plot(hist.history['accuracy'], label='Accuracy')
plt.plot(hist.history['val_accuracy'], label='Val Accuracy')
plt.plot(hist.history['loss'], label='Loss')
plt.plot(hist.history['val_loss'], label='Val Loss')
plt.legend()
plt.show()
```

N. Confusion Matrix

```
[ ] from sklearn.metrics import accuracy_score, confusion_matrix

matrix_cnn = confusion_matrix(y_test.argmax(axis=1), y_hat.argmax(axis=1))
print(matrix_cnn)
```

```
[ ] from sklearn.metrics import ConfusionMatrixDisplay
cmd = ConfusionMatrixDisplay(matrix_cnn,
                             display_labels=['Tumor', 'Stroma', 'Complex', 'Lympho',
                                             'Debris', 'Mucosa', 'Adipose', 'Empty'])

cmd.plot(cmap = plt.cm.Blues)
plt.title('Confusion Matrix')
```