

**KLASIFIKASI KANKER PARU BERDASARKAN CITRA HISTOPATOLOGI
DENGAN MENGGUNAKAN METODE CNN MODEL GOOGLNET**

PROPOSAL SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Oleh

Siti Nur Fadilah

NIM H92219061

PROGRAM STUDI MATEMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN AMPEL

SURABAYA

2022

DAFTAR ISI

| | |
|---|-------------------------------------|
| PERNYATAAN KEASLIAN | Error! Bookmark not defined. |
| LAPORAN PERSETUJUAN PEMBIMBING | Error! Bookmark not defined. |
| DAFTAR ISI | 2 |
| DAFTAR TABEL | 4 |
| DAFTAR GAMBAR..... | 5 |
| ABSTRAK..... | 6 |
| ABSTRACT | 7 |
| BAB I | 8 |
| PENDAHULUAN | 8 |
| 1.1 Latar Belakang Masalah..... | 8 |
| 1.2 Rumusan Masalah | 10 |
| 1.3 Tujuan Penelitian | 10 |
| 1.4 Manfaat Penelitian | 10 |
| 1.5 Sistematika Penulisan..... | 11 |
| BAB II..... | 12 |
| TINJAUAN PUSTAKA..... | 12 |
| 2.1 Kanker Paru | 12 |
| 2.2 Citra Digital | 13 |
| 2.2.1 Citra <i>Binary</i> | 13 |
| 2.2.2 Citra <i>Grayscale</i> | 14 |
| 2.2.3 Citra RGB (<i>Red Green Blue</i>)..... | 14 |
| 2.3 <i>Convolutional Neural Network (CNN)</i> | 15 |
| 2.3.1 <i>Convolution layer</i> | Error! Bookmark not defined. |
| 2.3.2 <i>Activation Function</i> | 16 |
| 2.3.3 <i>Pooling Layer</i> | 17 |
| 2.3.4 <i>Cross Channel Normalization</i> | 17 |

| | | |
|------------------------------------|------------------------------------|-------------------------------------|
| 2.3.5 | <i>Dropout</i> | Error! Bookmark not defined. |
| 2.3.6 | <i>Fully Connected Layer</i> | 18 |
| 2.3.7 | <i>Softmax Layer</i> | 18 |
| 2.4 | GoogleNet..... | 18 |
| 2.5 | <i>Confusion matrix</i> | 19 |
| 2.6 | Integrasi Keislaman..... | 20 |
| BAB III | | 22 |
| METODOLOGI PENELITIAN | | 22 |
| 3.1 | Jenis Penelitian..... | 22 |
| 3.2 | Sumber Data | 22 |
| 3.3 | .Kerangka Penelitian | 22 |
| DAFTAR PUSTAKA | | 24 |

DAFTAR TABEL

DAFTAR GAMBAR

ABSTRAK

KLASIFIKASI KANKER PARU BERDASARKAN CITRA HISTOPATOLOGI DENGAN MENGGUNAKAN METODE CNN MODEL GOOGLNET

Kata Kunci : GoogleNet, CNN, *Deep Learning*, Kanker paru

ABSTRACT

CLASSIFICATION OF LUNG CANCER BASED ON HISTOPATHOLOGICAL IMAGES USING THE CNN GOOGLNET MODEL

Kata Kunci : GoogleNet, CNN, *Deep Learning*, Lung Cancer

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Kanker merupakan salah satu masalah kesehatan global yang sangat penting karena merupakan penyakit nomor 3 yang dapat menyebabkan kematian setelah penyakit jantung (Mutiarasari 2019). Hal ini dikarenakan kanker adalah penyakit yang kompleks dan membutuhkan waktu yang relatif lama dalam proses pemulihannya. Tidak hanya waktu yang lama, proses pemulihan kanker juga tergolong sangat sulit. Kanker terjadi karena adanya proses mutasi sel yang kemudian menyebabkan sel tersebut tumbuh tanpa kendali dan melakukan pembelahan serta penyebaran yang sangat mudah dan cepat (Matthews, Bertoli, and de Bruin 2022). Penyebaran sel kanker yang cepat dan tidak terkontrol tersebut merupakan salah satu faktor utama mortalitas serta morbiditas di dunia. Terdapat banyak jenis kanker yang dapat terjadi seperti kanker payudara, kanker serviks, kanker paru, dan lain-lain (Shafabakhsh et al. 2019).

Kanker paru merupakan salah satu jenis kanker yang mengakibatkan kematian terbanyak. Di Indonesia, kanker paru menduduki peringkat ke-4 kematian terbanyak di dunia. Tingginya angka kematian akibat kanker paru salah satunya disebabkan karena adanya keterlambatan dalam mendiagnosis (ILLAH 2020). Hal ini dikarenakan sel kanker yang terdapat dalam tubuh sulit terdeteksi hingga akhirnya menyebar dan memasuki stadium tinggi. Pada kanker paru umumnya ditemukan adanya mutasi gen yang ditandai dengan perubahan urutan kromosom dalam sel (Wadowska et al. 2020). Banyak faktor yang dapat menyebabkan kanker paru seperti kebiasaan merokok, polusi, penyakit paru kronik, dan lain sebagainya. Berdasarkan faktor tersebut menyebabkan mengalami keluhan seperti batuk, nafas pendek, nyeri dada hingga batuk darah (YASMINE 2021).

Kanker paru secara umum dikelompokkan menjadi 2 kategori yaitu *Small Cell Lung Cancer* (SCLC) dan *Non-Small Cell Lung Cancer* (NSCLC) yang mana kedua kategori tersebut dibedakan berdasarkan pertumbuhan serta penyebarannya (Raso, Bota-Rabassedas, and Wistuba 2021). NSCLC merupakan kategori yang umum diderita dengan persentase 85% dari keseluruhan penderita kanker paru. Kategori NSCLC ini dibagi lagi menjadi 3 subkategori yakni *lung benign tissue*, *lung adenocarcinoma*, serta *lung squamous cell carcinoma* dimana dari ketiganya dibedakan berdasarkan letak sel

kankernya (Zulfa 2022). Terdapat beberapa pengobatan kanker yang harus dilakukan seperti kemoterapi, operasi, radioterapi, dan lain sebagainya. Namun, dari beberapa alternative pengobatan tersebut memiliki efek samping bagi tubuh serta cukup memakan waktu yang lama selama prosesnya. Oleh karena itu, penderita harus senantiasa bersabar sebagaimana firman Allah SWT:

يَا أَيُّهَا الَّذِينَ آمَنُوا اسْتَعِينُوا بِالصَّبْرِ وَالصَّلَاةِ إِنَّ اللَّهَ مَعَ الصَّابِرِينَ

Artinya: “Wahai orang-orang yang beriman! Mohonlah pertolongan (kepada Allah) dengan sabar dan salat. Sungguh, Allah beserta orang-orang yang sabar”.(QS. Al-Baqarah/2:153).

Berdasarkan ayat diatas, dijelaskan bahwa memohon pertolongan harus senantiasa untuk sabar dan salat karena Allah SWT selalu bersama orang-orang sabar. Jika senantiasa bersabar maka Allah SWT juga akan senantiasa memberikan pertolongan kepada hamba-Nya. Sekeras apapun sakit yang diderita jika Allah telah berkehendak untuk sembuh, maka akan hilanglah penyakit yang di deritanya. Hal ini selaras dengan firman Allah SWT:

إِنَّمَا أَمْرُهُ إِذَا أَرَادَ شَيْئًا أَنْ يَقُولَ لَهُ كُنْ فَيَكُونُ

Artinya: “Sesungguhnya urusan-Nya apabila Dia menghendaki sesuatu Dia hanya berkata kepadanya “Jadilah!” Maka jadilah sesuatu itu.”

Pemeriksaan untuk kanker paru dapat dilakukan dengan pemeriksaan laboratorium, radiologi, *trans thoracal biopsy* (TTB), *fine needle aspiration biopsy* (FNAB) serta pemeriksaan histopatogik (Joseph and Rotty 2020). Untuk pemeriksaan awal pendeteksian kanker paru bisa dilakukan dengan foto *thorax* AP. Jika terdapat tumor yang memiliki kemungkinan akan menjadi tumor ganas, maka dilanjutkan dengan pemeriksaan CT-Scan *thorax*. Pemeriksaan CT-Scan *thorax* ini juga dapat untuk mendiagnosa dan menentukan keganasan atas tumor tersebut (NURYANI n.d.). Di era perkembangan zaman ini, *Artificial Intelligence* (AI) mengembangkan teknologi yang dapat digunakan di bidang kedokteran salah satunya *Convolution Neural Network* (CNN).

CNN merupakan suatu metode gabungan antara jaringan saraf tiruan dengan deep learning yang umumnya digunakan untuk proses klasifikasi. Terdapat beberapa penelitian terdahulu yang menggunakan metode CNN yang memperoleh nilai akurasi yang tinggi seperti penelitian yang dilakukan oleh (Lestandy 2022) dengan memperoleh akurasi

sebesar 80%. Selain itu, juga terdapat penelitian yang dilakukan oleh (Fattah 2021) dengan nilai akurasi sebesar 97% dan penelitian oleh (Hasby 2022) dengan akurasi sebesar 100%.

Berdasarkan penelitian terdahulu, metode CNN sangat baik digunakan untuk klasifikasi karena akurasi yang diperoleh sangat tinggi. Terlebih untuk metode CNN model GoogleNet yang memiliki akurasi rata-rata diatas 90%. Oleh karena itu, pada penelitian ini digunakan metode CNN model GoogleNet untuk melakukan klasifikasi kanker paru berdasarkan citra histopatologi dengan menggunakan metode CNN model GoogleNet.

1.2 Rumusan Masalah

Berdasarkan masalah yang telah dipaparkan pada latar belakang, maka diperoleh rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana model optimal yang diperoleh pada sistem klasifikasi kanker paru dengan menggunakan metode CNN model GoogleNet berdasarkan uji coba parameter *Dropout* dan mini batchsize?
2. Bagaimana pengaruh augmentasi pada metode CNN model GoogleNet untuk diagnosis kanker paru?

1.3 Tujuan Penelitian

Berdasarkan dari beberapa rumusan masalah yang telah dipaparkan, maka tujuan yang ingin dicapai yaitu:

1. Memperoleh sistem diagnosis kanker paru yang terbaik dan tercepat dengan menggunakan metode CNN model GoogleNet berdasarkan hasil uji coba parameter *Dropout* dan mini batchsize.
2. Menganalisa pengaruh augmentasi pada metode CNN model GoogleNet untuk diagnosis kanker paru.

1.4 Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat kepada banyak pihak.

1. Manfaat Teoritis

a. Bagi Penulis

Manfaat bagi penulis yakni dapat digunakan sebagai salah satu proses pembelajaran terkait penulisan dan penyelesaian masalah secara sistematis serta dapat menambah keilmuan terkait kanker paru dan metode CNN model GoogleNet.

b. Bagi Pembaca

Manfaat bagi pembaca yakni dapat digunakan sebagai referensi keilmuan di bidang kedokteran dan menambah wawasan mengenai kanker paru dan metode CNN model GoogleNet.

2. Manfaat Praktis

Dalam dunia medis, penelitian ini digunakan oleh tenaga medis dalam membangun diagnosis kanker paru menggunakan metode CNN model GoogleNet.

1.5 Sistematika Penulisan

Sistematika pada penelitian ini dibagi menjadi beberapa bagian berikut:

1. BAB I PENDAHULUAN

Berisi latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

2. BAB II TINJAUAN PUSTAKA

Berisi landasan teori meliputi kanker, kanker paru, metode CNN, GoogleNet dan *confusion matrix*.

3. BAB III METODE PENELITIAN

Berisi jenis penelitian, jenis dan sumber data serta tahap-tahap penelitian yang dilakukan.

4. BAB IV HASIL DAN PEMBAHASAN

Berisi hasil dari penerapan metode CNN model GoogleNet dalam klasifikasi paru serta menganalisis hasil dari penerapan tersebut.

5. BAB V PENUTUP

Berisi kesimpulan dari hasil penelitian yang telah dilakukan dan saran untuk penelitian selanjutnya.

BAB II

TINJAUAN PUSTAKA

2.1 Kanker Paru

Kanker adalah salah satu penyakit yang sering berujung pada kematian. Berbagai macam jenis kanker yang dapat menyebabkan kematian. Terdapat salah satu jenis kanker mencapai dengan persentase hampir 13% menyebabkan kematian di dunia yaitu kanker paru (LALANG, SILALAH, and BUKHARI 2018). Di Indonesia, WHO menyatakan bahwa penderita kanker paru terbanyak yaitu berjenis kelamin laki-laki. Kanker paru merupakan neoplasma ganas yang berasal dari epitel bronkus, dimana keganasan dari kanker paru tersebut dapat berasal dari paru itu sendiri (primer) ataupun dari luar paru (metastatis). Kanker paru adalah salah satu jenis kanker yang berasal dari sel epitelium respirasi. Berdasarkan jenisnya sel tersebut terbagi menjadi dua yaitu *Small Cell Lung Cancer* (SCLC) dan *Non-Small Cell Lung Cancer* (NSCLC) (Hynds et al. 2021). *Small Cell Lung Cancer* (SCLC) merupakan jenis sel kecil yang diisi mukus dan berkumpul diantara pembuluh darah halus. Sedangkan, *Non-Small Cell Lung Cancer* (NSCLC) masih terbagi lagi menjadi tiga subjenis yaitu *adenocarcinoma* yang menempati 35% hingga 40% paru, *squamous cell carcinoma* yang memiliki ciri khas adanya pembentukan jembatan intraseluler, dan *large cell carcinoma*. Jenis NSCLC merupakan jenis kanker paru yang sering dijumpai dimana hampir semua kasus kanker paru 70% hingga 80% adalah penderita NSCLC dengan berbagai macam faktor penyebab (Patel 2020).

Faktor penyebab kanker paru dapat berasal dari polusi dan kebiasaan merokok. Perokok pasif juga dapat terkena kanker paru yang disebabkan oleh paparan asap rokok dari perokok aktif. Sedangkan, bagi perokok aktif mempunyai nilai kemungkinan terkena kanker paru lebih besar dibandingkan dengan perokok pasif (Satriawan 2022). Selain asap rokok, gas radon juga dapat menyebabkan terkena kanker paru dikarenakan gas tersebut berasal dari pemecahan uranium. Umumnya, tidak banyak penderita kanker disebabkan karena gas tersebut namun jika berada dalam suatu ruangan dan terdapat gas radon itu dapat meningkatkan nilai persentase kemungkinan terkena kanker paru (Buana and Harahap 2022). Penyebab terkena kanker paru tidak hanya itu saja, namun adapula faktor umum seperti perubahan genetik sesuai gender dan umur, pola hidup tidak sehat, dan riwayat keluarga. Selain memperhatikan penyebab terkena kanker paru, gejala awal penderita paru

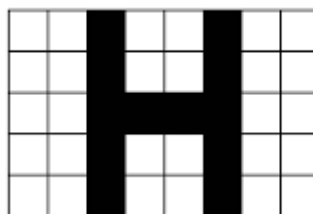
juga merupakan hal penting yang harus diperhatikan. Gejala penderita kanker paru diantaranya batuk berdarah, berat badan yang mengalami penurunan drastik, dan yang paling banyak dirasakan yaitu sesak nafas (NIRMALA 2020).

2.2 Citra Digital

Citra adalah sebuah gambar dua dimensi yang didapatkan dari gambar analog dua dimensi yang kemudian gambar tersebut dibagi menjadi N baris dan M kolom sehingga menghasilkan gambar diskrit. Proses pada citra tersebut dinamakan dengan proses sampling (Sembiring, Efendi, and Suwilo 2020). Pada proses ini dibagi menjadi dua bagian yaitu proses *upsampling* dimana pada proses ini dapat meningkatkan resolusi dari suatu gambar dengan menaikkan jumlah piksel. Sedangkan, pada proses lainnya yaitu *downsampling* untuk memperoleh resolusi citra spasial terhadap suatu gambar dengan menurunkan jumlah piksel. Citra didefinisikan dengan fungsi $f(x,y)$ dimana x dan y merupakan koordinat spasial dari sebuah citra, serta $f(x,y)$ merupakan intensitas dari citra pada titik tersebut (PRASETIO 2021). Citra digital merupakan sebuah gambar dua dimensi yang dapat diolah dengan menggunakan *computer* dan menghasilkan sebuah nilai. Nilai-nilai tersebut merupakan nilai intensitas pada tiap masing-masing piksel. Berdasarkan jenis warna, citra digital dibagi menjadi beberapa macam diantaranya:

2.2.1 Citra *Binary*

Citra *binary* atau biner merupakan suatu citra yang hanya memiliki dua kemungkinan yaitu hitam dan putih pada setiap pikselnya. Citra ini juga dapat disebut sebagai citra monokrom dimana pada citra ini hanya membutuhkan 1 bit yang cukup untuk mewakili nilai setiap pikselnya sama halnya dengan citra *grayscale* (Saputra, Yudhana, and Umar 2022). Umumnya, citra ini didapatkan melalui proses pengolahan citra seperti segmentasi yang kemudian diproses ulang untuk mendeteksi lokasi suatu objek dengan nilai piksel kecil. Berikut contoh citra *binary* ditunjukkan oleh Gambar 2.1.



Gambar 2.1 Contoh Citra *Binary*

2.2.2 Citra *Grayscale*

Citra *grayscale* atau citra beraras keabuan adalah suatu citra digital dengan tingkat intensitas nilai *Red*, *Green*, dan *Blue* nya sama. Citra ini hanya memiliki warna hitam, keabuan, serta putih. Skala keabuan pada citra ini berada pada interval 0 hingga 255 dimana semakin tinggi nilai skala keabuan suatu citra maka semakin terang pula derajat keabuannya (Hsu et al. 2021). Umumnya, citra ini direpresentasikan ke dalam bentuk array dua dimensi dimana tiap elemen yang terdapat pada array tersebut menunjukkan intensitas dari suatu citra yang berada pada koordinat yang bersesuaian dengan jumlah bit yang digunakan adalah 8 bit. Berikut contoh citra *grayscale* ditunjukkan oleh Gambar 2.2.



Gambar 2.2 Contoh Citra *Grayscale*

2.2.3 Citra RGB (*Red Green Blue*)

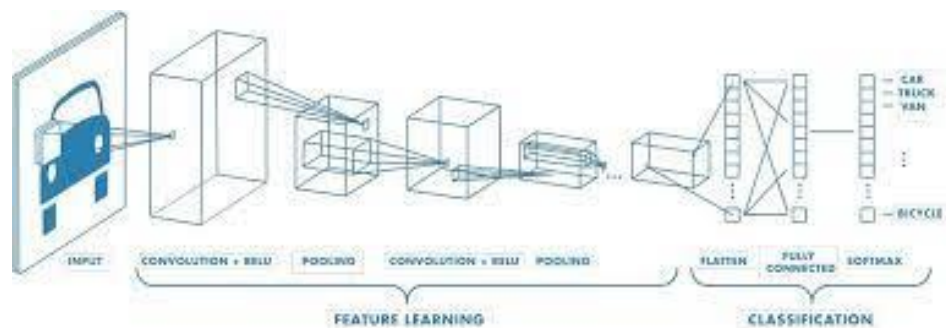
Citra RGB atau juga dikenal dengan citra warna merupakan suatu citra yang terdiri dari tiga warna dasar yakni merah, hijau dan biru yang jika digabungkan akan menghasilkan banyak warna. Setiap warna dasar pada citra ini masing-masing memiliki intensitas sebesar 8 bit yakni berada pada interval 0 hingga 255 (Theilmann 2022). Jika setiap warna dasar memiliki nilai sebesar 255 maka memerlukan media untuk penyimpanan sebesar 3 byte untuk tiap piksel. Selain itu, juga akan mendapatkan sebanyak 16 juta jenis warna yang diperoleh dari masing-masing nilai maksimum setiap warna dasar. Berikut contoh citra RGB ditunjukkan oleh Gambar 2.3.



Gambar 2.3 Contoh Citra RGB

2.3 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan salah satu jenis *deep learning* yang dirancang untuk mengolah data dua dimensi karena memiliki hasil yang sangat signifikan dalam pengenalan citra (Ghulam et al. 2022). Oleh karena itu, CNN dirancang khusus untuk mengklasifikasikan suatu citra karena kedalaman jaringan pada CNN tergolong tinggi dan berusaha untuk mencontoh sistem pengenalan citra yang terdapat pada *visual cortex* manusia. Sama seperti metode *deep learning* lainnya, CNN juga memiliki kelemahan yaitu membutuhkan waktu yang lama untuk proses pelatihan model (Wu et al. 2018). Namun, kini kelemahan tersebut dapat teratasi dengan menggunakan teknologi *General Purpose Graphical Processing Unit (GPGPU)*. Pada metode ini terdapat beberapa komponen utama seperti *Convolution layer*, *Pooling layer*, *Dropout layer*, dan lain sebagainya dimana komponen-komponen tersebut ditunjukkan oleh Gambar 2.4 yang merupakan arsitektur dari CNN.



Gambar 2.4 Arsitektur CNN

Berdasarkan gambar diatas, terdapat beberapa komponen utama pada arsitektur CNN diantaranya:

2.3.1 Convolution layer

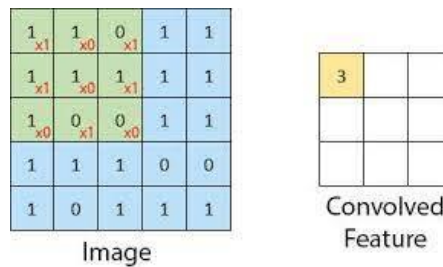
Convolution layer adalah *layer* pertama yang berada pada CNN dimana *layer* ini bertugas sebagai penerima input citra dan melakukan ekstraksi fitur pada citra tersebut. Komponen ini merupakan komponen utama yang mendasar yang berada pada CNN dimana seluruh data yang ada akan mengalami proses konvolusi. Umumnya, pada *Convolution layer* ini dilakukan suatu operasi yakni operasi kombinasi linear filter terhadap daerah lokal dengan filter yang ukuran panjang, lebar, dan tinggi sesuai dengan data citra yang diterima (Kartikasari 2020). Rumus yang digunakan pada operasi konvolusi ditunjukkan pada persamaan 2.1.

$$Q_{(i,j)} = \left(\sum_{i=1}^N I_{(i,j)} \times F_{(i,j)} \right) + B_{(i,j)} \quad (2.1)$$

Berdasarkan rumus di atas diketahui bahwa $Q(i, j)$ merupakan hasil perkalian antara matriks input yaitu $I(i, j)$ dengan matriks filter yaitu $F(i, j)$ yang kemudian ditambahkan dengan nilai bias. Terdapat beberapa parameter yang digunakan pada proses ini, yaitu *stride* dan *padding* (Sadewa, Irawan, and Setianingsih 2019). *Stride* merupakan suatu parameter yang tujuannya untuk menentukan jumlah pergeseran pada suatu filter, dimana semakin kecil *stride* komputasi yang dibutuhkan lebih besar. *Padding* atau juga dikenal dengan *zero padding* merupakan suatu parameter yang digunakan untuk menentukan jumlah pixel yang bernilai 0 serta mengatur dimensi output tetap sama agar dapat dilakukan proses ekstraksi fitur yang mendalam. Perhitungan untuk memperoleh dimensi output dari *feature map* ditunjukkan oleh persamaan 2.2.

$$H = \frac{I - F + 2p}{s} + 1 \quad (2.2)$$

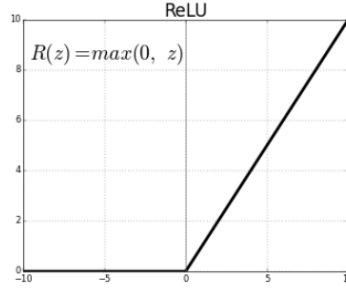
Berdasarkan rumus di atas diketahui bahwa H adalah output dari *feature map*, I merupakan matriks input, F merupakan matriks filter, p merupakan *padding*, dan s merupakan *stride*. Ilustrasi proses pada *convolution layer* ditunjukkan oleh gambar 2.5.



Gambar 2.5 Ilustrasi *Convolution layer*

2.3.2 Activation Function

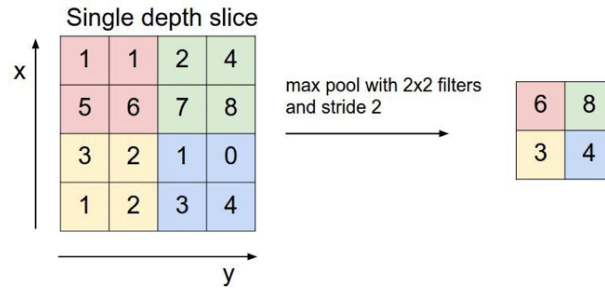
Fungsi aktivasi merupakan suatu fungsi non-linier yang digunakan untuk menyelesaikan permasalahan non trivial. Terdapat beberapa macam fungsi aktivasi, salah satunya fungsi aktivasi ReLU. *Rectified Linear Unit* atau ReLU merupakan salah satu fungsi aktivasi untuk menghilangkan nilai negatif yang ada pada citra dengan cara menggantinya dengan nilai 0 (Haq 2021). Fungsi aktivasi ReLU ini dapat membantu untuk mengurangi waktu *training* dan testing.



Gambar 2.6 Ilustrasi Activation Function

2.3.3 Pooling Layer

Pooling layer merupakan suatu *layer* yang terdiri dari sebuah filter yang memiliki ukuran dan *stride* tertentu yang nantinya akan bergeser di area *feature map* secara bergantian serta untuk mempercepat komputasi karena telah dimensi *feature map* telah berkurang (Lee and Song 2019). Umumnya, terdapat dua macam yang sering digunakan yaitu *max pooling* yang mengambil nilai paling maksimal dan *average pooling* yang mengambil nilai rata-rata.



Gambar 2.7 Ilustrasi Pooling Layer

2.3.4 Cross Channel Normalization

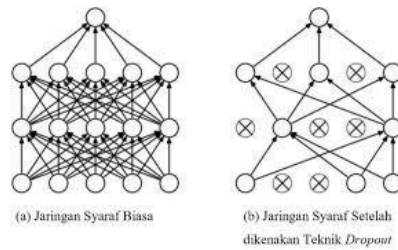
Cross channel normalization digunakan untuk meningkatkan kemampuan *generalisasi* dimana mengubah nilai setiap elemen dengan nilai normalisasi yang berdasarkan nilai elemen dari fitur terdekat baik dari *feature map* yang sama ataupun tidak (Phiadelvira 2021). Perhitungan yang digunakan ditunjukkan oleh persamaan 2.3.

$$x_{i,j}^p = \frac{a_{i,j}^p}{\left(k + \frac{\alpha \sum_{\max(1, p-\frac{n}{2})}^{\min(N-1, p+\frac{n}{2})} (a_{i,j}^p)^2}{n} \right)^\beta} \quad (2.3)$$

Dimana $x_{i,j}^p$ merupakan output dari *feature map*, $a_{i,j}^p$ merupakan *feature map* serta beberapa konstanta yang digunakan seperti k , α , dan β .

2.3.5 Dropout

Dropout merupakan suatu teknik dalam algoritma jaringan saraf yang digunakan untuk mengantisipasi adanya *overfitting* dengan menggabungkan beberapa jenis arsitektur. Jika semakin tinggi nilai bobot maka akan proses *training* akan membutuhkan waktu yang semakin lama (Virkhansa, Setiawan, and Dewi 2019). Oleh karena itu, jika semakin tinggi tingkat *Dropout* yang digunakan, maka semakin *general* pula model yang diperoleh.



Gambar 2.8 Ilustrasi *Dropout Layer*

2.3.6 Fully Connected Layer

Pada *layer* ini setiap neuron yang memiliki keterkaitan hingga menjadi multi *layer* perceptron. *Layer* ini merupakan *layer* terakhir setelah *convolutional layer* dan *Pooling layer* yang digunakan sebagai *layer* penyatuan akhir yang melakukan klasifikasi ke dalam kelas pada data *training* (Nasrulloh 2020). *Fully connected layer* ini didefinisikan oleh persamaan 2.4.

$$y_j = b_j + \sum_i w_{ij}x_i \quad (2.4)$$

Berdasarkan persamaan diatas dapat diketahui bahwa y_j merupakan output dari *fully connected layer*, b_j merupakan bias, w_{ij} merupakan bobot, dan x_i merupakan input yang didapat dari learning feature.

2.3.7 Softmax Layer

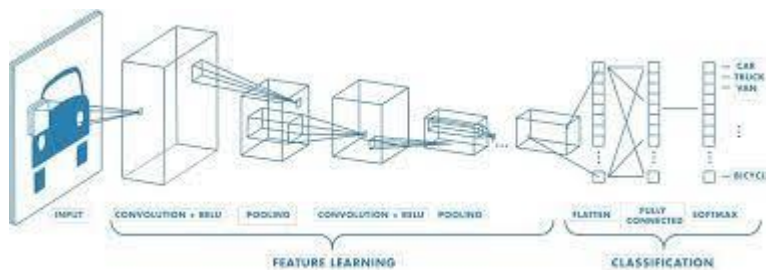
Softmax merupakan salah satu fungsi aktivasi yang digunakan untuk mengubah nilai vector serta untuk memperoleh probabilitas tiap data dengan interval 0 hingga 1. Perhitungan *softmax layer* ini ditunjukkan oleh persamaan 2.5.

$$p_j(x) = \frac{e^{x_j}}{\sum_{i=1}^k e^{x_i}} \quad (2.5)$$

2.4 GoogleNet

GoogleNet merupakan salah satu model CNN yang pertama kali dikenalkan oleh Google pada tahun 2014. CNN model ini memiliki inception modules yang terdiri atas

sejumlah convolution kecil yang berguna untuk mereduksi (Fattah 2021). Akurasi atas model ini akan tinggi jika kinerja pada jaringan sarafnya meningkat yaitu dengan memperdalam *layer*. Algoritma CNN model GoogleNet memiliki 144 *layer* dimana terdiri atas operasi konvolusi, *Pooling*, batch normalization serta klasifikasi (Jawahar, Sharen, and Gandomi 2022). Ilustrasi arsitektur pada CNN model GoogleNet ditunjukkan oleh Gambar 2.9.



Gambar 2.9 Arsitektur GoogleNet

2.5 Confusion matrix

Proses evaluasi dalam melakukan klasifikasi suatu hal merupakan salah satu proses penting (Normawati and Prayogi 2021). Umumnya, proses ini menggunakan *confusion matrix* yaitu suatu matriks yang menyatakan hasil klasifikasi pada data yang teruji benar dan data yang teruji salah. Selain itu, *confusion matrix* juga dapat digunakan untuk menentukan nilai persentase keakuratan, sensitivitas, serta spesifitas suatu model dengan menggunakan 4 istilah yang mempresentasikan hasil klasifikasinya (ENDAH and ENCIS n.d.). Istilah-istilah pada *confusion matrix* yaitu *True Positive* (TP) dimana sejumlah data yang terklasifikasi benar pada kelas aktual dan kelas prediksi, *True Negative* (TN) dimana sejumlah data yang terklasifikasi salah pada kelas aktual dan kelas prediksi, *False Positive* (FP) dimana sejumlah data yang terklasifikasi salah pada kelas aktual dan benar pada kelas prediksi, dan *False Negative* (FN) dimana sejumlah data yang terklasifikasi benar pada kelas aktual dan salah pada kelas prediksi (Chicco, Tötsch, and Jurman 2021). Pada penelitian ini, menggunakan *confusion matrix multi class* dimana terdapat kelas *benign lung tissue*, *lung adenocarcinoma*, dan *lung squamous cell carcinoma* yang dimisalkan dengan kelas X,Y, dan Z secara berturut-turut. Tabel *confusion matrix multi class* ditunjukkan pada Gambar 2.10.

| Kelas Aktual | Kelas klasifikasi | | |
|--------------|-------------------|----|----|
| | A | B | C |
| A | TP | FP | FP |
| B | FP | TP | FP |
| C | FP | FP | TP |

| Kelas Aktual | Kelas klasifikasi | | |
|--------------|-------------------|----|----|
| | A | B | C |
| A | TP | FN | FN |
| B | FN | TP | FN |
| C | FN | FN | TP |

| Kelas Aktual | Kelas klasifikasi | | |
|--------------|-------------------|----|----|
| | A | B | C |
| A | TN | TP | TN |
| B | TP | TP | TP |
| C | TN | TP | TN |

Gambar 2.10 *Confusion Matrix Multi Class*

Berdasarkan gambar diatas, maka dapat dilakukan perhitungan nilai akurasi, sensitivitas, serta spesifitas atas hasil klasifikasi. Akurasi digunakan untuk mengukur tingkat keakuratan suatu system. Sensitivitas digunakan untuk memperoleh jumlah data yang terklasifikasi positif. Sedangkan, spesifisitas digunakan digunakan untuk memperoleh jumlah data yang terklasifikasi positif pada kelas negatif. Perhitungan nilai akurasi, sensitivitas, serta spesifisitas ditunjukkan oleh persamaan 2.5.

$$\begin{aligned}
 \text{Akurasi} &= \frac{TP_{all}}{n_{all}} \times 100\% \\
 \text{Sensitivitas} &= \frac{\sum \frac{TP}{TP+FN}}{n} \times 100\% \\
 \text{Spesifitas} &= \frac{\sum \frac{TN}{TN+FN}}{n} \times 100\%
 \end{aligned}
 \tag{2.5}$$

2.6 Integrasi Keislaman

Penelitian ini yang mana mengklasifikasikan tipe jenis kanker paru diharapkan bermanfaat bagi semua masyarakat di semua bidang. Hal ini dikarenakan jika kita memberikan manfaat bagi orang lain, maka kita sendiri juga akan manfaat atas hal tersebut. Hal ini seperti pada firman Allah SWT:

...إِنْ أَحْسَنْتُمْ أَحْسَنْتُمْ لِأَنْفُسِكُمْ

Artinya : “Jika kalian berbuat baik, sesungguhnya kalian berbuat baik bagi diri kalian sendiri ...” (QS al-isra/17:7).

Manusia di dunia tidak lepas dari suatu ujian, dimana ujian tersebut merupakan suatu tolak ukur bagi manusia atas keimanannya kepada Allah SWT. Ujian yang diturunkan dapat berupa ujian ni'mat (kesenangan) maupun niqmat (kesengsaraan). Kesenangan sama halnya dengan ujian karena dibalik kesenangan tersirat yakni ujian. Kesengsaraan merupakan suatu ujian yang juga tidak dapat dihindari. Hal ini selaras dengan kisah Nabi Ayyub yang diuji dalam bentuk ni'mat yaitu dengan diberikan kekayaan harta serta niqmat yaitu berupa kemiskinan serta sakit. Seperti pada firman Allah SWT:

أَحْسِبَ النَّاسَ أَنْ يُلْزَمُوا أَنْ يَقُولُوا ءَامَنَّا وَهُمْ لَا يُفْتَنُونَ
وَلَقَدْ فَتَنَّا الَّذِينَ مِنْ قَبْلِهِمْ فَلَيَعْلَمَنَّ اللَّهُ الَّذِينَ صَدَقُوا وَلَيَعْلَمَنَّ الْكَاذِبِينَ

Artinya: “Apakah manusia itu mengira bahwa mereka dibiarkan (saja) mengatakan “Kami telah beriman”, sedang mereka tidak diuji lagi? Dan sesungguhnya kami telah menguji orang-orang yang sebelum mereka, maka sesungguhnya Allah mengetahui orang-orang yang benar dan sesungguhnya Dia mengetahui orang-orang yang dusta.”(QS. Al-Ankabut/29:2-3)

مَا أَصَابَ مِنْ مُصِيبَةٍ إِلَّا بِإِذْنِ اللَّهِ وَمَنْ يُؤْمِنْ بِاللَّهِ يَهْدِ اللَّهُ قَلْبَهُ وَاللَّهُ بِكُلِّ شَيْءٍ عَلِيمٌ

Ujian bentuk niqmat dapat berupa penyakit baik siapa saja dan dimana saja mendapatkan ujian tersebut. Namun, kita sebagai manusia harus bisa menerima dengan lapang dada dan menjalaninya. Hal ini selaras dengan kewajiban orang yang sakit dimana mereka harus ikhlas atas apa yang terjadi padanya. Sebagaimana dijelaskan dalam firman Allah SWT:

Artinya: “Tidak ada sesuatu musibah yang menimpa (seseorang), kecuali dengan izin Allah; dan barangsiapa beriman kepada Allah, niscaya Allah akan memberi petunjuk kepada hatinya. Dan Allah Maha Mengetahui segala sesuatu”.(QS. At Taghobun/64:11).

BAB III

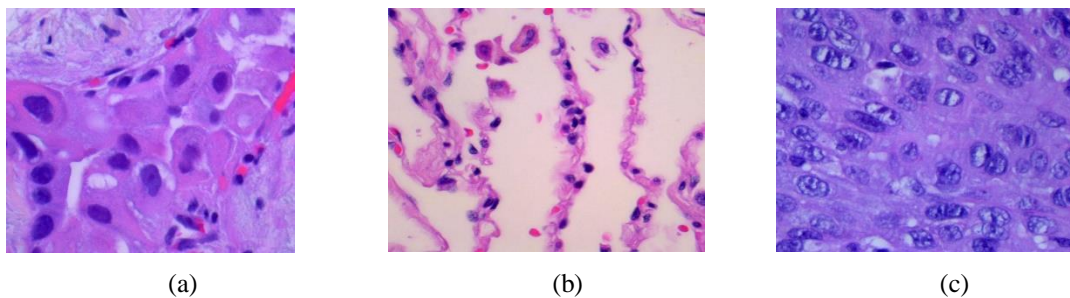
METODOLOGI PENELITIAN

3.1 Jenis Penelitian

Penelitian ini dikategorikan ke dalam jenis penelitian kuantitatif. Pada penelitian ini data yang digunakan merupakan data citra histopatologi kanker paru. Berdasarkan data tersebut kemudian dilakukan pengolahan sehingga menghasilkan sebuah matriks, dimana pada matriks tersebut berisi nilai-nilai piksel yang ada pada citra tersebut. Kemudian, dilakukan proses perhitungan dengan menggunakan matriks tersebut hingga memperoleh hasil diagnose pada masing-masing citra.

3.2 Sumber Data

Pada penelitian ini, data yang digunakan diperoleh melalui website Kaggle.com. Data yang diperoleh berjumlah 15000 data yang terbagi menjadi 3 kelas yaitu *benign lung tissue*, *lung adenocarcinoma*, dan *lung squamous cell carcinoma* dimana masing-masing kelas berisi 5000 data. Sampel data citra kanker paru ditunjukkan oleh Gambar 3.1.

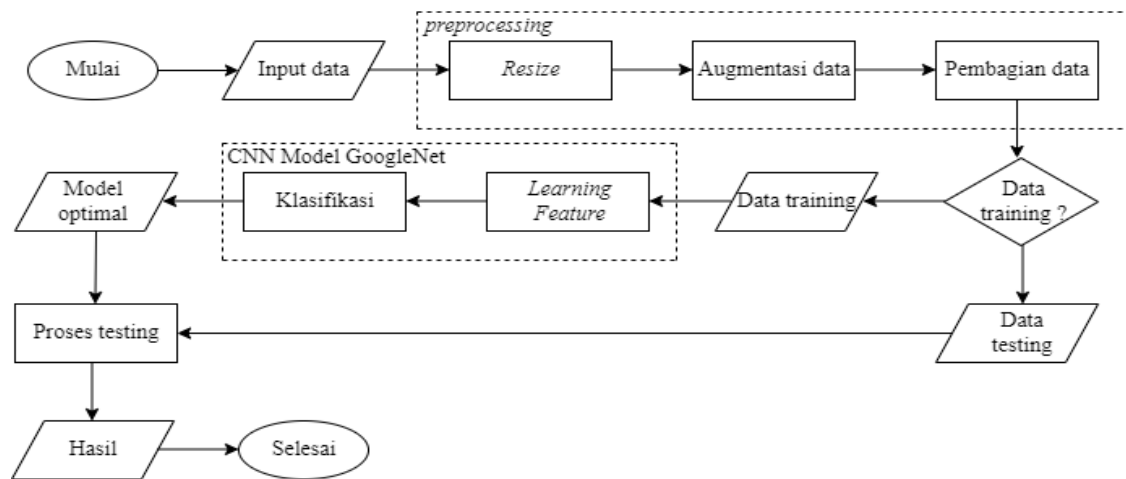


Gambar 3.1 Citra Histopatologi Kanker Paru

(a)*Lung benign tissue*; **(b)** *Lung adenocarcinoma*; **(c)** *Lung squamous cell carcinoma*

3.3. Kerangka Penelitian

Pada penelitian ini terdapat beberapa proses penelitian dalam melakukan klasifikasi kanker paru dengan menggunakan CNN model GoogleNet. Proses yang dilakukan ditunjukkan oleh Gambar 3.2.



Gambar 3.2 Alur Penelitian

Pada penelitian ini terdapat beberapa tahapan yang dilakukan, berikut tahapannya yaitu:

1. Tahap preprocessing dilakukan proses resizing pada data citra histopatologi kanker paru yang semula berukuran 768 x 768 piksel menjadi 224 x 224 menyesuaikan dengan ukuran inputan data pada model GoogleNet. Kemudian dilanjutkan dengan proses augmentasi data dimana pada proses ini dilakukan refleksi serta rotasi. Kemudian, pembagian data dengan SplitEachLabel dengan beberapa persentase data *training* yang kemudian akan diuji coba dengan system.
2. Tahap klasifikasi dengan CNN model GoogleNet yaitu terdapat proses learning feature. Proses ini dilakukan dengan menerapkan arsitektur model GoogleNet dimana terdapat *Convolution layer*, activation function, *Pooling layer*, cross channel normalization, dan fully connected *layer*.
3. Tahap selanjutnya yaitu dilakukan uji coba parameter dengan beberapa variasi *Dropout*. Setelah itu, dilanjutkan dengan proses klasifikasi.
4. Tahap selanjutnya yaitu evaluasi hasil dengan menggunakan confusion matrix untuk memperoleh nilai akurasi, sensitivitas, dan spesifitas suatu model.

DAFTAR PUSTAKA

- Buana, Indra, and Dwi Agustian Harahap. 2022. "ASBESTOS, RADON DAN POLUSI UDARA SEBAGAI FAKTOR RESIKO KANKER PARU PADA PEREMPUAN BUKAN PEROKOK." *AVERROUS: Jurnal Kedokteran Dan Kesehatan Malikussaleh* 8(1):1–16.
- Chicco, Davide, Niklas Tötsch, and Giuseppe Jurman. 2021. "The Matthews Correlation Coefficient (MCC) Is More Reliable than Balanced Accuracy, Bookmaker Informedness, and Markedness in Two-Class Confusion Matrix Evaluation." *BioData Mining* 14(1):1–22.
- ENDAH, FAUZININGRUM, and INDAH SURYANINGSIIH ENCIS. n.d. "EVALUASI DAN PREDIKSI PENGUASAAN BAHASA INGGRIS MARITIM MENGGUNAKAN METODE DECISION TREE DAN CONFUSION MATRIX (STUDI KASUS DI UNIVERSITAS MARITIM AMNI)." *PROSIDING KEMARITIMAN 2021*.
- Fattah, Muhammad Syaifulloh. 2021. "Deteksi Penyakit Pneumonia Dan COVID-19 Menggunakan Citra x-Ray Dengan Metode Convolutional Neural Network (CNN) Model GoogleNet."
- Ghulam, Ali, Farman Ali, Rahu Sikander, Ashfaq Ahmad, Aftab Ahmed, and Shruti Patil. 2022. "ACP-2DCNN: Deep Learning-Based Model for Improving Prediction of Anticancer Peptides Using Two-Dimensional Convolutional Neural Network." *Chemometrics and Intelligent Laboratory Systems* 226:104589.
- Haq, Dina Zatusiva. 2021. "Klasifikasi Citra Kanker Kulit Menggunakan Convolutional Neural Network Model Googlenet."
- Hasby, Alfi Nur Fitri Ali. 2022. "Sistem Diagnosa Citra Histopatologi Kanker Ginjal Menggunakan Metode Cnn Model Googlenet."
- Hsu, Chen-Ming, Chien-Chang Hsu, Zhe-Ming Hsu, Feng-Yu Shih, Meng-Lin Chang, and Tsung-Hsing Chen. 2021. "Colorectal Polyp Image Detection and Classification through Grayscale Images and Deep Learning." *Sensors* 21(18):5995.
- Hynds, Robert E., Kristopher K. Frese, David R. Pearce, Eva Grönroos, Caroline Dive, and Charles Swanton. 2021. "Progress towards Non-Small-Cell Lung Cancer Models That

- Represent Clinical Evolutionary Trajectories.” *Open Biology* 11(1):200247.
- ILLAHI, ANNISA N. U. R. 2020. “Analisis Faktor Yang Berhubungan Dengan Perilaku Pencarian Pengobatan Alternatif Pada Pasien Kanker Di Yayasan Kanker Indonesia Cabang Jawa Timur.”
- Jawahar, Malathy, H. Sharen, and Amir H. Gandomi. 2022. “ALNett: A Cluster Layer Deep Convolutional Neural Network for Acute Lymphoblastic Leukemia Classification.” *Computers in Biology and Medicine* 148:105894.
- Joseph, Junita, and Linda W. A. Rotty. 2020. “Kanker Paru: Laporan Kasus.” *Medical Scope Journal (MSJ)* 2(1).
- Kartikasari, Mujiati Dwi. 2020. “Implementasi Deep Learning Object Detection Rambu K3 Pada Video Menggunakan Metode Convolutional Neural Network (CNN) Dengan Tensorflow (Studi Kasus: Rambu Kesehatan Dan Keselamatan Kerja (K3) Jalur Evakuasi Dan Alat Pemadam Api Pada Gedung FMIPA UII).”
- LALANG, D., B. P. SILALAH, and F. BUKHARI. 2018. “Vehicle Routing Problem Time Windows Dengan Pengemudi Sese kali.” *Journal of Mathematics and Its Applications* 17(2):87–99. doi: 10.29244/jmap.17.2.87-99.
- Lee, Hagyeong, and Jongwoo Song. 2019. “Introduction to Convolutional Neural Network Using Keras; an Understanding from a Statistician.” *Communications for Statistical Applications and Methods* 26(6):591–610.
- Lestandy, Merinda. 2022. “Deteksi Dini Kanker Payudara Menggunakan Metode Convolution Neural Network (CNN).” *Inspiration: Jurnal Teknologi Informasi Dan Komunikasi* 12(1):65–72.
- Matthews, Helen K., Cosetta Bertoli, and Robertus A. M. de Bruin. 2022. “Cell Cycle Control in Cancer.” *Nature Reviews Molecular Cell Biology* 23(1):74–88.
- Mutiarasari, Diah. 2019. “Ischemic Stroke: Symptoms, Risk Factors, and Prevention.” *Medika Tadulako: Jurnal Ilmiah Kedokteran Fakultas Kedokteran Dan Ilmu Kesehatan* 6(1):60–73.
- Nasrulloh, Nur Zam Zam. 2020. “TA: PERBANDINGAN LAYER ARSITEKTUR RESIDUAL NETWORK UNTUK DETEKSI PENGGUNAAN SABUK PENGAMAN PADA PENGEMUDI MOBIL.”

- NIRMALA, SANTI GITA. 2020. "ASUHAN KEPERAWATAN DENGAN MASALAH KETIDAKEFEKTIFAN BERSIHAN JALAN NAFAS PADA PASIEN TUBERCULOSIS PARU DIRSU DR. WAHIDIN SUDIRO HUSODO MOJOKERTO."
- Normawati, Dwi, and Surya Allit Prayogi. 2021. "Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter." *J-SAKTI (Jurnal Sains Komputer Dan Informatika)* 5(2):697–711.
- NURYANI, SISKI. n.d. "Klasifikasi Nodul Paru Pada Citra Ct Scan Berdasarkan Ciri Tekstur Dan Ciri Morfologi Menggunakan Jaringan Saraf Tiruan Propagasi Balik Dengan Scilab Dan Weka."
- Patel, Vivek Navinchandra. 2020. "Delivery of SiRNA for Masking Resistance to Chemotherapy in Non-Small Cell Lung Cancer."
- Phiadelvira, Bunga Yuwa. 2021. "Klasifikasi Kanker Serviks Berdasarkan Citra Kolposkopi Menggunakan Convolutional Neural Network (CNN) Model Alexnet."
- PRASETIO, ANNAS. 2021. "CITRA DIGITAL DAN ALGORITMA PENERAPANNYA." Thesis Commons.
- Raso, Maria Gabriela, Neus Bota-Rabassedas, and Ignacio I. Wistuba. 2021. "Pathology and Classification of SCLC." *Cancers* 13(4):820.
- Sadewa, Raam Pujangga, Budhi Irawan, and Casi Setianingsih. 2019. "Fire Detection Using Image Processing Techniques with Convolutional Neural Networks." Pp. 290–95 in *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*. IEEE.
- Saputra, Sabarudin, Anton Yudhana, and Rusydi Umar. 2022. "Implementation of Naïve Bayes for Fish Freshness Identification Based on Image Processing." *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)* 6(3):412–20.
- Satriawan, Dodi. 2022. "Gambaran Kebiasaan Merokok Penduduk Di Indonesia." *Jurnal Litbang Sukowati: Media Penelitian Dan Pengembangan* 5(2):51–58.
- Sembiring, Rafika Sari, Syahril Efendi, and Saib Suwilo. 2020. "Improving the Accuracy of Old and Young Face Detection in the Template Matching Method with Fuzzy Associative Memory (FAM)." P. 12117 in *IOP Conference Series: Materials Science and Engineering*. Vol. 725. IOP Publishing.

- Shafabakhsh, Rana, Russel J. Reiter, Hamed Mirzaei, Somayyeh Noei Teymoordash, and Zatollah Asemi. 2019. "Melatonin: A New Inhibitor Agent for Cervical Cancer Treatment." *Journal of Cellular Physiology* 234(12):21670–82.
- Theilmann, Florian. 2022. "Hands-on-Experiments for Additive Colour Mixing." *Physics Education* 57(5):55031.
- Virkhansa, Chelsa Farah, Budi Darma Setiawan, and Candra Dewi. 2019. "Klasifikasi Status Gunung Berapi Dengan Metode Learning Vector Quantization (LVQ)." *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer E-ISSN 2548:964X*.
- Wadowska, Katarzyna, Iwona Bil-Lula, Łukasz Trembecki, and Mariola Śliwińska-Mossoń. 2020. "Genetic Markers in Lung Cancer Diagnosis: A Review." *International Journal of Molecular Sciences* 21(13):4569.
- Wu, Haifeng, Qing Huang, Daqing Wang, and Lifu Gao. 2018. "A CNN-SVM Combined Model for Pattern Recognition of Knee Motion Using Mechanomyography Signals." *Journal of Electromyography and Kinesiology* 42:136–42.
- YASMINE, ULAYYA CHAIRUNISSA. 2021. "KARAKTERISTIK KLINIS DAN PATOLOGIS PASIEN KANKER PARU DENGAN SINDROMA VENA KAVA SUPERIOR SEBUAH KAJIAN LITERATUR."
- Zulfa, Irkhana Indaka. 2022. "Klasifikasi Kanker Paru Berdasarkan Citra Histopatologi Menggunakan Metode Convolutional Neural Network (CNN) Model AlexNet."