

**EVALUASI KINERJA BERBAGAI METODE *MACHINE LEARNING*
PADA KLASIFIKASI KANKER KULIT**

(Skripsi)

Oleh

**SAFIIRA RAHMAH LINISA
2017051015**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024**

**EVALUASI KINERJA BERBAGAI METODE *MACHINE LEARNING*
PADA KLASIFIKASI KANKER KULIT**

Oleh

**SAFIIRA RAHMAH LINISA
201701015**

Skripsi

**Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA KOMPUTER**

Pada

**Jurusan Ilmu Komputer
Fakultas Matematika dan Ilmu Pengetahuan Alam**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2024**

ABSTRAK

EVALUASI KINERJA BERBAGAI METODE *MACHINE LEARNING* PADA KLASIFIKASI KANKER KULIT

Oleh

SAFIIRA RAHMAH LINISA

Kanker kulit telah menjadi perhatian yang signifikan dan terus meningkat dalam beberapa tahun terakhir, terutama karena menipisnya lapisan ozon dan meningkatnya paparan radiasi ultraviolet yang berbahaya. Deteksi dini kanker sangat penting untuk pengobatan yang tepat waktu dan efektif, yang secara signifikan meningkatkan peluang kesembuhan. Di tengah perkembangan teknologi, penelitian ini memperkenalkan pendekatan *machine learning* untuk klasifikasi kanker kulit ganas dan jinak. Penelitian ini menggunakan beberapa arsitektur *Convolutional Neural Network* diantaranya *Custom Arsitektur*, Resnet50, DenseNet121 dan DenseNet169. Selain itu, menggunakan model *Support Vector Machine* (SVM) dengan beberapa kernel, yang menggunakan teknik ekstraksi fitur bentuk, tekstur, dan warna. Di antara model CNN, DenseNet-169, dengan *optimizer* SGD, mencapai akurasi rata rata tertinggi sebesar 0,8908, secara signifikan mengungguli arsitektur lainnya. Dalam SVM, *kernel linear* menghasilkan kinerja terbaik, dengan akurasi 0,8300. Penggunaan *dataset* yang seimbang dan *k-vold cross validation* pada model DenseNet-169 menunjukkan efektivitas yang baik dalam klasifikasi kanker kulit.

Kata kunci : Machine Learning, CNN, SVM, Kanker Kulit

ABSTRACT

PERFORMANCE EVALUATION OF VARIOUS MACHINE LEARNING METHODS FOR SKIN CANCER CLASSIFICATION

By

SAFIIRA RAHMAH LINISA

Skin cancer has become a significant and growing concern in recent years, primarily due to the depletion of the ozone layer and increased exposure to harmful ultraviolet radiation. Early detection is crucial for timely and effective treatment, which greatly improves the chances of a cure. Leveraging advancements in technology, this study presents machine learning approaches for classifying malignant and benign skin cancers. It employs several Convolutional Neural Network (CNN) architectures, including Custom Architecture, ResNet-50, DenseNet-121, and DenseNet-169. Additionally, a Support Vector Machine (SVM) model with various kernels is utilized, incorporating shape, texture, and color feature extraction techniques. Among the CNN models, DenseNet-169, optimized with SGD, achieved the highest average accuracy of 0.8908, significantly outperforming the other architectures. For the SVM models, the linear kernel demonstrated the best performance, with an accuracy of 0.8300. The use of balanced datasets and k-fold cross-validation with the DenseNet-169 model proved to be highly effective for skin cancer classification.

Keywords: Machine Learning, CNN, SVM, Skin Cancer.

Judul Skripsi : **EVALUASI KINERJA BERBAGAI
METODE *MACHINE LEARNING* PADA
KLASIFIKASI KANKER KULIT**

Nama Mahasiswa : **Safiira Rahmah Tinisa**

Nomor Pokok Mahasiswa : 2017051015

Program Studi : S1 Ilmu Komputer

Jurusan : Ilmu Komputer

Fakultas : Matematika dan Ilmu Pengetahuan Alam



A handwritten signature in black ink, consisting of a large loop at the top and a vertical line extending downwards.

Prof. Admi Syarif, Ph.D.
NIP. 19670103199203 1 003

A handwritten signature in blue ink, featuring a large, stylized initial 'D' and 'S'.

Dewi Asiah Shofiana, S.Komp., M.Kom.
NIP. 19950929202012 2 030

2. Ketua Jurusan Ilmu Komputer

A handwritten signature in blue ink, consisting of several overlapping loops and lines.

Dwi Sakethi, S.Si., M.Kom
NIP. 19680611 199802 1 001

MENGESAHKAN

1. Tim Penguji

Ketua : Prof. Admi Syarif, Ph.D.

Sekretaris : Dewi Asiah Shofiana S.Komp., M.Kom.

Penguji Utama : Dr. rer. nat. Akmal Junaidi, M.Sc.



Three handwritten signatures in blue ink are positioned to the right of the text. The top signature is the longest and most stylized, corresponding to the Chairman. The middle signature is shorter and more compact, corresponding to the Secretary. The bottom signature is also shorter and more compact, corresponding to the Main Examiner.

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Heri Satria, S.Si., M.Si.
NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi : 23 Oktober 2024

PERNYATAAN

Saya yang bertanda tangan di bawah ini, menyatakan bahwa skripsi saya yang berjudul “**Evaluasi Kinerja Berbagai Metode *Machine Learning* Pada Klasifikasi Kanker Kulit**“ merupakan karya saya sendiri dan bukan karya orang lain. Semua tulisan yang tertuang di skripsi ini telah mengikuti kaidah penulisan karya tulis ilmiah Universitas Lampung. Apabila di kemudian hari terbukti skripsi saya merupakan hasil jiplakan atau dibuat orang lain, maka saya bersedia menerima sanksi berupa pencabutan gelar yang saya terima.

Bandar Lampung, 23 Oktober 2024

Penulis,



Safiira Rahmah Linisa

NPM. 2017051015

RIWAYAT HIDUP



Penulis lahir di Bandar Lampung pada tanggal 22 Mei 2002 sebagai anak kedua dari pasangan Bapak M.Isyrofi dan Ibu Siti Wahyuni. Penulis masuk Pendidikan formal di SD Negeri 1 Poncowarno pada tahun 2014, kemudian melanjutkan di SMP Negeri 1 Kalirejo dan selesai pada tahun 2017. Penulis kemudian melanjutkan pendidikan menengah atas di SMA Negeri 1 Kalirejo dan lulus pada tahun 2020.

Pada tahun 2020 penulis terdaftar sebagai mahasiswa Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur undangan atau SNMPTN. Selama menjadi mahasiswa penulis aktif mengikuti berbagai kegiatan di antaranya sebagai berikut.

1. Menjadi anggota pengurus Unit Kegiatan Mahasiswa Rohani Islam (Rois) Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung tahun periode 2021 dan 2022.
2. Menjadi Asisten dosen untuk mata kuliah matematika pada semester ganjil tahun ajaran 2021/2022.
3. Mengikuti Kursus Pemrograman Berorientasi Objek pada Program Kredensial Mikro Mahasiswa Indonesia (KMMI) tahun 2021.
4. Mengikuti program Kampus Merdeka Studi Independen Bersertifikat Batch 3 di Dicoding Academy alur pembelajaran *Machine Learning* dan *Front-End Web* tahun 2022.

5. Menjadi Bendahara Dinas Pemberdayaan Wanita Badan Eksekutif Mahasiswa (BEM) Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung tahun 2023.
6. Melaksanakan KKN Mandiri di Desa Fajar Baru, Kecamatan Jati Agung, Kabupaten Lampung Selatan pada tahun 2023.
7. Mengikuti program Bangkit Academy Kampus Merdeka jalur pembelajaran *Machine Learning* pada tahun 2023 Batch 1.
8. Menjadi penerima pendanaan Program Mahasiswa Wirausaha (PMW) Universitas Lampung tahun 2023.

MOTTO

“Maka sesungguhnya bersama kesulitan ada kemudahan. Sesungguhnya bersama kesulitan ada kemudahan.”

(QS. Al-Insyirah [94]: 5-6)

“Ingatlah, Sesungguhnya pertolongan Allah itu dekat.”

(QS. Al-Baqarah[2]: 214)

“Jihad paling utama adalah jihad melawan nafsu sendiri karena Allah.”

- Nabi Muhammad Salallahu Alaihi Wasalam

(HR. Ahmad, At-Tirmidzi, Abu Dawud, dan Shahih Ibn Hibban)

“Jangan bilang tidak mungkin kepadaku sebelum kamu mati mencobanya.”

- Sultan Muhammad Al-Fatih

(Sang Penakluk Konstantinopel)

PERSEMBAHAN

Alhamdulillahirobbilalamin

Dengan rasa syukur kepada Allah Subhanahu Wa Ta'ala atas Ridho, Rahmat, dan Karunia-Nya sehingga saya dapat menyelesaikan skripsi ini. Shalawat serta salam selalu tercurahkan kepada Nabi Muhammad Shallallahu Alaihi Wasallam.

Kupersembahkan karya ini kepada:

Kedua Orang Tuaku Tercinta

Ibu dan Ayah, terima kasih selalu atas doa dan kasih sayang yang tiada terhingga. Terima kasih atas segala dukungan baik moril ataupun materi. Terima kasih atas doa yang tidak pernah lelah diberikan, terima kasih atas segala keringat dan air mata. Terima kasih untuk selalu percaya. Terima kasih untuk kakak dan adikku tersayang atas dukungan dan do'a yang diberikan kepadaku.

Terima kasih juga untuk semua pihak yang terlibat dalam pembuatan skripsi ini. Semoga Allah senantiasa membalas setiap kebaikan kalian.

Almamater Tercinta, Universitas Lampung dan Jurusan Ilmu Komputer

Tempat bernaung mengemban ilmu untuk bekal kehidupan.

SANWACANA

Puji syukur penulis ucapkan kehadirat Allah Subhanahu Wata'ala, karena atas rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini. Shalawat serta salam semoga senantiasa tercurahkan kepada junjungan Nabi Muhammad Shallallahu Alaihi Wasallam yang penulis harapkan syafaatnya di hari akhir kelak. Skripsi dengan judul "**Evaluasi Kinerja Berbagai Metode *Machine Learning* Pada Klasifikasi Kanker Kulit**" adalah salah satu syarat untuk memperoleh gelar sarjana komputer pada Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

Dalam penyusunan skripsi ini, penulis mendapat bantuan, dukungan dan motivasi dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar besarnya kepada:

1. Kedua orang tuaku tercinta, Ayah dan Ibu yang selalu mendoakan, mendukung, memotivasi dan memberikan kasih sayang pada penulis.
2. Bapak Prof. Admi Syarif, Ph.D., selaku dosen pembimbing utama yang selalu membimbing, memberikan arahan, motivasi, masukan serta saran, dalam penyelesaian skripsi ini.
3. Ibu Dewi Asiah Shofiana, S.Komp., M.Kom., selaku dosen pembimbing kedua yang selalu membimbing, memberikan arahan, masukan dan saran dalam penyelesaian skripsi ini.
4. Bapak Dr. rer. nat. Akmal Junaidi, M.Sc., selaku dosen pembahas yang telah memberikan saran dan masukan yang bermanfaat dalam perbaikan skripsi ini.
5. Bapak Dwi Sakethi, S.Si., M.Kom. selaku Ketua Jurusan Ilmu Komputer FMIPA Universitas Lampung.

6. Ibu Anie Rose Irawati, S.T., M.Cs. selaku Sekretaris Jurusan Ilmu Komputer FMIPA Universitas Lampung.
7. Bapak, Ibu Dosen dan semua Staf Jurusan Ilmu Komputer FMIPA Universitas Lampung yang telah memberikan banyak ilmu dan wawasan baru kepada penulis dan telah membantu segala urusan administrasi yang penulis butuhkan.
8. Kakak dan Adikku yang selalu mendukung, memberikan do'a dan semangat kepada penulis.
9. Teman seperjuangan semasa kuliah, Nafasya, Hida, Nuk, Aniisah, Irma, Silvia, Putri, Yulia dan teman teman lainnya yang tidak bisa penulis sebutkan satu persatu, terima kasih karena selalu ada, terima kasih telah mendukung, mendoakan, membantu, dan mewarnai hari penulis selama masa perkuliahan.
10. Lela dan Pepy yang selalu menemani, mendukung dan menjadi keluarga penulis selama tinggal di bandar lampung.
11. Teman teman dan Ibu Bapak seperbimbingan, Bu Apri, Bu Lilis, Pak Nurjoko, Pak Rico, Pak Nizar, Joy, Arman, Yoan dan Nabila yang telah memberikan bantuan, masukan, dukungan, dan menemani selama pengerjaan skripsi semoga Allah selalu memudahkan.
12. Teman-teman Ilmu Komputer angkatan 2020 yang menjadi teman satu angkatan selama menjalankan masa studi.

Akhir kata, penulis menyadari bahwa skripsi ini masih jauh dari kesempurnaan, akan tetapi penulis berharap skripsi ini dapat berguna dan bermanfaat bagi semua. Semoga dengan bantuan dan dukungan yang diberikan mendapat balasan pahala di sisi Allah Subhanahu Wa Ta'ala. Aamiin.

Bandar Lampung, 23 Oktober 2024

Penulis,



Safiira Rahmah Linisa

NPM. 2017051015

DAFTAR ISI

	Halaman
DAFTAR ISI.....	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL.....	xvii
I. PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	5
II. TINJAUAN PUSTAKA	6
2.1 Penelitian Terdahulu.....	6
2.2 <i>Machine Learning</i>	8
2.3 <i>Support Vector Machine</i>	8
2.4 <i>Deep Learning</i>	12
2.5 <i>Convolutional Neural Network</i>	13
2.5.1 <i>Layer CNN</i>	14
2.5.2 <i>Rectified Linear Unit (ReLU)</i>	17
2.5.3 <i>Batch Normalization</i>	18
2.5.4 <i>Flatten</i>	18
2.5.5 <i>Dropout</i>	18
2.6 <i>Image Augmentation</i>	19
2.7 <i>Confusion Matrix</i>	21
2.8 <i>K-Fold Cross Validation</i>	23
2.9 Kanker Kulit	24
III. METODE PENELITIAN.....	27

3.1	Tempat dan Waktu	27
3.1.1	Tempat penelitian.....	27
3.1.2	Waktu Penelitian	27
3.2	Alat dan Bahan	29
3.2.1	<i>Hardware</i>	29
3.2.2	<i>Software</i>	29
3.2.3	<i>Library</i>	29
3.3	Tahap Penelitian	31
3.3.1	Akuisisi Data.....	31
3.3.2	<i>Pre-Processing</i>	32
3.3.3	SVM	33
3.3.4	CNN	34
3.3.5	Evaluasi	35
IV.	HASIL DAN PEMBAHASAN.....	37
4.1	Akuisisi Data	37
4.2	<i>Data Pre-processing</i>	38
4.2.1	<i>Balancing Data</i>	38
4.2.2	Normalisasi data.....	41
4.2.3	Pembagian data	41
4.2.1	Ekstraksi Fitur SVM	42
4.3	Pemodelan SVM.....	46
4.4	Pemodelan CNN.....	51
4.5	Perbandingan Algoritma.....	72
V.	PENUTUP.....	76
5.1	Simpulan.....	76
5.2	Saran.....	77
	DAFTAR PUSTAKA	78

DAFTAR GAMBAR

Gambar	Halaman
1. Ilustrasi metode SVM (Kurniadi et al., 2021).....	9
2. Contoh jaringan <i>deep learning</i> (Nahr et al., 2021).	13
3. <i>Convolutional Neural Network</i> (Yamashita et al., 2018).....	14
4. <i>Convolutional Layer</i> (Yamashita et al., 2018).	15
5. <i>Average Pooling vs Max Pooling</i> (Agarwal & Singh, 2022).....	16
6. <i>Dropout Layer</i> (Chen et al., 2018).	19
7. Proses <i>k-fold cross validation</i> (Lyu et al., 2022).	24
8. (a) Kanker kulit ganas, (b) Kanker kulit jinak (www.kaggle.com).	26
9. Alur kerja penelitian.....	31
10. Kanker kulit jinak (www.kaggle.com).....	32
11. Kanker kulit ganas (www.kaggle.com).....	32
12. Contoh hasil teknik rotasi.....	40
13. Contoh hasil teknik <i>zoom</i>	40
14. Contoh hasil teknik <i>horizontal flip</i>	41
15. Contoh hasil ekstraksi fitur warna.....	43
16. Contoh hasil ekstraksi fitur bentuk.	44
17. Ekstraksi fitur tekstur dengan LBP.	44
18. Hasil ekstraksi fitur teksur GLCM.	45
19. <i>Confusion matrix linear kernel</i> skema 1 48	48
20. <i>Confusion matrix</i> skema 2 <i>linear kernel</i> 50	50
21. Grafik metrik evaluasi model SVM 51	51
22. Arsitektur <i>Custom Model</i> 53	53
23. Ilustrasi Arsitektur DenseNet (Huang et al., 2016)..... 53	53
24. Arsitektur DenseNet (Huang et al., 2016)..... 55	55
25. Ilustrasi arsitektur ResNet (Gomes et al., 2022). 56	56
26. <i>Layer</i> arsitektur Resnet (He et al., 2016). 56	56
27. Grafik perbandingan evaluasi arsitektur CNN skema 1..... 58	58
28. <i>Confusion matrix</i> arsitektur DenseNet169 skema 1..... 59	59
29. Metrik evaluasi perbandingan arsitektur CNN skema 2. 60	60
30. <i>Confusion matrix</i> DenseNet169 skema 2..... 61	61

31. Perbandingan metrik evaluasi pada skema 1.	63
32. Perbandingan metrik evaluasi pada skema 2.	64
33. Metrik evaluasi <i>optimizer</i> dan <i>learning rate</i> skema 1.....	66
34. Metrik evaluasi <i>optimizer</i> dan <i>learning rate</i> skema 2.....	68
35. Perbandingan <i>epoch</i> skema 1	70
36. Metrik evaluasi <i>epoch</i> skema 2.....	71
37. Grafik perbandingan metrik evaluasi SVM dan CNN.	73
38. Grafik perbandingan akurasi algoritma <i>machine learning</i>	74

DAFTAR TABEL

Tabel	Halaman
1. Penelitian terdahulu.....	6
2. Jadwal penelitian.....	28
3. Skenario pembagian <i>dataset</i>	33
4. Distribusi kelas.....	38
5. Data skema 1	39
6. Data skema 2.....	39
7. Jumlah data berdasarkan skenario pembagian data	42
8. Hasil akurasi SVM dengan nilai $k = 5$	46
9. Hasil akurasi SVM dengan nilai $k = 10$	47
10. Hasil akurasi model SVM berbagai kernel pada skema 1.....	48
11. Hasil akurasi model SVM berbagai kernel pada skema 2.....	49
12. Metrik evaluasi SVM skema 1 dan skema 2.....	50
13. <i>Hyperparameter</i> CNN.....	51
14. Perbandingan hasil akurasi arsitektur CNN skema 1.	57
15. Perbandingan metrik evaluasi arsitektur CNN Skema 1.....	57
16. Perbandingan hasil akurasi arsitektur CNN skema 2.	59
17. Metrik evaluasi arsitektur CNN skema 2.	60
18. Hasil akurasi perbandingan <i>batch size</i> pada skema 1.	62
19. Metrik evaluasi perbandingan <i>batch size</i> skema 1.....	63
20. Hasil akurasi perbandingan <i>batch size</i> pada skema 2.	63
21. Metrik evaluasi <i>batch size</i> skema 2.....	64
22. Hasil akurasi perbandingan <i>learning rate</i> pada skema 1 data.....	65
23. Metrik evaluasi <i>optimizer dan learning rate</i> skema 1.	66
24. Perbandingan hasil akurasi <i>optimizer learning rate</i> pada skema 2.	67
25. Metrik evaluasi <i>optimizer dan learning rate</i> skema 2.....	68
26. Hasil akurasi perbandingan <i>epoch</i> pada skema 1.....	69
27. Metrik evaluasi perbandingan <i>epoch</i> skema 1.	69
28. Hasil akurasi perbandingan <i>epoch</i> pada skema 2.....	70
29. Metrik evaluasi perbandingan <i>epoch</i> skema 2.	71
30. Metrik evaluasi SVM dan CNN	72
31. Perbandingan akurasi algoritme <i>machine learning</i>	74

I. PENDAHULUAN

1.1 Latar Belakang

Kanker kulit menjadi salah satu kanker yang paling luas tersebar di dunia (Aishwarya et al., 2023). Kanker kulit telah menjadi masalah serius dalam beberapa tahun terakhir, terutama disebabkan oleh penipisan lapisan ozon di langit dan meningkatnya paparan radiasi ultraviolet (Aishwarya et al., 2023). Kanker kulit terbentuk ketika terjadi kesalahan (mutasi) DNA sel-sel kulit yang mengakibatkan pertumbuhan sel di luar kendali dan membentuk massa sel kanker (Agarwal & Singh, 2022). Beberapa gejala awal yang berkaitan dengan kanker ini meliputi perubahan warna atau peradangan pada kulit, sensasi gatal dan pendarahan pada bercak kulit, tahi lalat atau bercak merah, serta munculnya benjolan seperti lilin pada kulit karena pertumbuhan sel kanker (Agarwal & Singh, 2022).

Tumor jinak dapat termasuk jenis kanker dan umum dianggap sebagai tahi lalat, yang tidak berbahaya (Monika et al., 2020). Tumor ganas dikenal sebagai kanker yang dapat mengancam kehidupan, serta dapat merusak jaringan tubuh lainnya (Monika et al., 2020). Lesi kulit awalnya dapat berupa tahi lalat, plak, atau papul (Oktora et al., 2023). Jika terdapat tahi lalat yang tumbuh dengan cepat, memiliki ukuran tidak teratur, terasa nyeri, dan mudah berdarah, maka dapat dicurigai sebagai tumor ganas (Oktora et al., 2023). Deteksi dini kanker sangatlah penting untuk penanganan dan penyembuhan penyakit sejak awal. Hal ini berlaku untuk kanker kulit yang dapat diidentifikasi, diobati, dan disembuhkan seperti penyakit lainnya (Khater et al., 2023).

Dalam bidang kesehatan diagnosis digunakan untuk mendeteksi adanya kelainan. Dokter biasanya melakukan diagnosis kanker kulit dengan menggunakan biopsi. Kelemahan dari proses biopsi ini adalah biayanya dibutuhkan cukup banyak, dapat melukai kulit manusia sebagai sampel dan prosesnya tidak cukup efektif (Khasanah et al., 2021). Prosesnya yaitu dengan mengambil sampel kecil dari jaringan kulit kemudian melakukan pemeriksaan di laboratorium (Faruk & Nafi'iyah, 2020). Perlu dikembangkan teknologi alternatif untuk mendiagnosis penyakit kanker kulit, dengan harapan bahwa dengan adanya teknologi tersebut, kanker kulit dapat terdeteksi lebih cepat dan akurat.

Seiring dengan pesatnya perkembangan teknologi, *machine learning* semakin banyak digunakan dalam bidang kesehatan. *Machine learning* dapat mempermudah proses klasifikasi penyakit seperti mengetahui jenis penyakit dan memberikan hasil dalam bentuk gambar yang lebih optimal (Telaumbanua et al., 2019). *Deep learning* merupakan bagian dari pengembangan *machine learning*. *Deep learning* menghilangkan kebutuhan terhadap rekayasa fitur karena dapat secara otomatis mengekstraksi fitur-fitur penting dari data mentah (Ameri, 2020). Pada penelitian sebelumnya *machine learning* dan *deep learning* telah dimanfaatkan dalam menyelesaikan masalah pengklasifikasian kanker kulit.

Penelitian (Faruk & Nafi'iyah, 2020) menggunakan *dataset Skin Cancer: Malignant vs. Benign* dari Kaggle *dataset* dengan jumlah sebanyak 3297 citra, yang terdiri dari 2637 citra sebagai *data training*, dan 660 citra sebagai *data testing*. Penelitian ini mengklasifikasikan kanker kulit jinak dan ganas dan diperoleh hasil akurasi tertinggi sebesar 70,61% pada algoritme KNN dengan $K=5$, nilai akurasi SVM 69,85% dan dengan *Naive Bayes* hanya 67,27%. Perbandingan antara metode SVM, *Random Forest* dan *K-Nearest Neighbor* untuk deteksi kanker kulit juga dilakukan pada penelitian (Murugan et al., 2019) dengan data dari *International Skin Imaging Collaboration (ISIC)* sebanyak 1000 data. Metode SVM menghasilkan akurasi tertinggi untuk *ABCD rule* sebesar 89.43%, kemudian *Random*

Forest 76.87% dan KNN menghasilkan akurasi sebesar 69.54% (Murugan et al., 2019).

Pada penelitian (Savera et al., 2020) dengan jumlah *dataset* sebanyak 800 data yang terdiri dari 400 data kanker kulit *malignant* dan 400 data *benign* klasifikasi menggunakan metode KNN menghasilkan akurasi sebesar 75% sementara klasifikasi menggunakan metode CNN menghasilkan akurasi sebesar 76%. Pada penelitian (Agarwal & Singh, 2022) *dataset testing* berisi 350 citra sedangkan *dataset testing* berisi lebih dari 2900 citra yang terdiri dari dua kelas yang berbeda, yaitu "jinak" dan "ganas". Dengan percobaan menggunakan beberapa arsitektur seperti *DenseNet*, *Resnet*, *XceptionNet*, dan *MobileNet* diperoleh hasil akurasi tertinggi sebesar 86% menggunakan arsitektur *Resnet* (Agarwal & Singh, 2022).

Dari beberapa penelitian yang ada salah satu algoritme *machine learning* yang banyak digunakan untuk kasus klasifikasi adalah *Support Vector Machine*. *Support Vector Machine* (SVM) telah terbukti efektif dalam menangani *dataset* berdimensi tinggi dan relatif kecil, yang sesuai dengan karakteristik *dataset* kanker kulit. Sementara itu, *Convolutional Neural Network* (CNN) memiliki kemampuan mengekstraksi fitur dari suatu gambar yang sesuai digunakan untuk citra medis.

Pada penelitian ini akan dibangun sebuah model *Convolutional Neural Network* dan *Support Vector Machine* kemudian mengevaluasi kinerja kedua metode tersebut dan membandingkan dengan metode yang digunakan pada penelitian sebelumnya. Evaluasi kedua metode tersebut digunakan untuk mendapatkan model dengan akurasi terbaik dan menghasilkan diagnosis yang lebih akurat. *Dataset* diperoleh dari kaggle *dataset* yang merupakan arsip ISIC (*International Skin Imaging Collaboration*). *Dataset* terdiri dari 3297 data yang terdiri dari dua kelas yang berbeda "jinak" dan "ganas".

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang, rumusan masalah dari penelitian ini adalah sebagai berikut.

- a. Bagaimana membangun model klasifikasi kanker kulit jinak dan ganas menggunakan metode *Convolutional Neural Network*.
- b. Bagaimana membangun model klasifikasi kanker kulit dengan menggunakan metode *Support Vector Machine*.
- c. Bagaimana performa metode *Convolutional Neural Network* dan *Support Vector Machine* dibandingkan dengan metode *machine learning* lain pada penelitian terdahulu dalam mengklasifikasikan kanker kulit.

1.3 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut.

- a. Membangun model klasifikasi kanker kulit dengan menggunakan metode *Convolutional Neural Network*.
- b. Membangun model klasifikasi kanker kulit dengan menggunakan metode *Support Vector Machine*.
- c. Evaluasi model *Convolutional Neural Network* dan *Support Vector Machine* serta membandingkan performa akurasi dengan metode pada penelitian terdahulu agar mendapatkan metode yang paling efektif.

1.4 Manfaat Penelitian

Manfaat yang akan diperoleh dari penelitian ini antara lain sebagai berikut.

- a. Mengetahui hasil kinerja dari hasil dari *Convolutional Neural Network* dan *Support Vector Machine* dan perbandingan dengan metode pada penelitian sebelumnya untuk klasifikasi kanker kulit.
- b. Menjadi alternatif untuk diagnosis dini penyakit kanker kulit sehingga pengobatan lebih awal dapat dilakukan dan dapat meningkatkan peluang sembuh bagi pasien.
- c. Hasil penelitian ini dapat memberikan informasi yang berguna untuk memahami dan mengembangkan teknologi klasifikasi kanker kulit yang lebih baik.
- d. Penelitian ini dapat digunakan untuk pengembangan penelitian selanjutnya.

1.5 Batasan Masalah

Adapun batasan masalah dalam penelitian ini adalah sebagai berikut.

- a. Metode klasifikasi kanker kulit ganas dan jinak menggunakan *Convolutional Neural Network* dan *Support Vector Machine*.
- b. Menggunakan *dataset* dari Kaggle yang terdiri dari dua jenis citra tahi lalat jinak (*benign skin moles*) dan tahi lalat ganas (*malignant skin moles*).
- c. Menggunakan metode *Image Augmentation* pada data untuk menangani kasus *imbalanced data*.

II. TINJAUAN PUSTAKA

2.1 Penelitian Terdahulu

Terdapat beberapa penelitian terdahulu yang berhubungan dengan penelitian ini yang diuraikan dalam Tabel 1.

Tabel 1. Penelitian terdahulu

Penelitian	Metode	Data	Hasil Akurasi
<i>Classification of Skin Cancer Images using Convolutional Neural Networks</i> (Agarwal & Singh, 2022).	CNN dengan arsitektur <i>DenseNet</i> , <i>Resnet</i> , dan <i>MobileNet</i> .	<i>Kaggle dataset</i> , <i>malignant vs benign</i> yang berisi 2900 citra gambar.	<i>DenseNet</i> 86%, <i>Resnet</i> 86,571%, <i>XceptionNet</i> 82,571% dan <i>MobileNet</i> 80,857%.
Klasifikasi Kanker Kulit Berdasarkan Fitur Tekstur, Fitur Warna Citra Menggunakan SVM dan KNN (Faruk & Nafi'iyah, 2020).	Algoritme SVM, KNN, dan <i>Naive Bayes</i> .	<i>Skin Cancer: Malignant vs. Benign</i> dari <i>Kaggle dataset</i> dengan jumlah sebanyak 3297 citra gambar.	SVM 69,85%, KNN dengan nilai K=5 akurasi 70,61%, dan <i>Naive Bayes</i> 67,27%.

Tabel 1. (Lanjutan)

Penelitian	Metode	Data	Hasil Akurasi
<i>Detection of Skin Cancer Using SVM, Random Forest and K-NN Classifiers</i> (Murugan et al., 2019).	SVM, <i>Random Forest</i> dan K-NN.	<i>International Skin Imaging Collaboration</i> (ISIC) sebanyak 1000 data.	SVM 89.43%, <i>Random Forest</i> 76.87% dan KNN 69.54%.
Deteksi Dini Kanker Kulit menggunakan K-NN dan <i>Convolutional Neural Network</i> (Savera et al., 2020).	K-NN dan <i>Convolutional Neural Network</i> .	800 data yang terdiri dari 400 data kanker kulit <i>malignant</i> dan 400 data <i>benign</i> .	KNN 75%, CNN 76%.
Klasifikasi Kanker Kulit Menggunakan Algoritme <i>Random Forest</i> (Khasanah et al., 2021).	<i>Random Forest Algorithm</i> .	<i>Skin Cancer: Malignant vs. Benign</i> dari Kaggle <i>dataset</i> dengan jumlah sebanyak 3297 citra gambar.	Akurasi 84,25%.

2.2 *Machine Learning*

Machine learning adalah subbidang *Artificial Intelligence* yang memungkinkan sistem komputer belajar dari data dan membuat prediksi tentang data yang akan datang dengan menggunakan pengetahuan yang diperoleh melalui proses pembelajaran tersebut (Ayyıldız, 2023). Kata “*learning*” tidak hanya memiliki arti pengetahuan atau pemahaman melalui membaca, pendidikan maupun pengalaman tetapi juga meningkatkan performanya melalui pengalaman (Ayyıldız, 2023). *Machine learning* menghidupkan mesin dengan kemampuan belajar secara mandiri dengan kata lain, mesin tersebut dapat belajar melalui pengalaman, pengamatan, dan analisis pola berdasarkan data yang tersedia (Nahr et al., 2021).

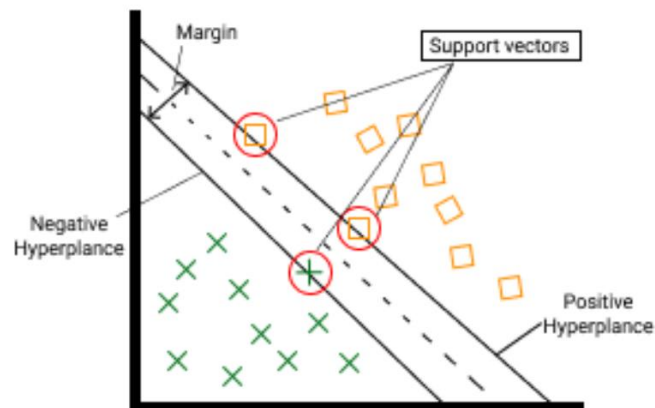
Penerapan *machine learning* telah menghasilkan dampak yang cukup signifikan pada bidang kesehatan (Javaid et al., 2022). *Machine learning* memiliki tujuan untuk meningkatkan kecepatan dan keakuratan tugas seorang dokter dengan membantu dalam mendeteksi tanda-tanda awal epidemi atau pandemi (Javaid et al., 2022). Algoritme ini dapat menganalisis data dari berbagai sumber, seperti memeriksa data satelit, laporan berita dan media sosial, dan bahkan sumber video untuk menentukan apakah penyakit akan menjadi tidak terkendali (Javaid et al., 2022). Penggunaan *machine learning* dalam perawatan kesehatan dapat membuka peluang yang memungkinkan penyedia layanan kesehatan untuk lebih fokus pada perawatan pasien daripada pada pengumpulan atau pengecekan informasi (Javaid et al., 2022).

2.3 *Support Vector Machine*

Support Vector Machine (SVM) merupakan metode yang sangat populer dalam bidang pemrosesan data dan *machine learning*. SVM digunakan untuk mengklasifikasikan informasi dalam bentuk linier dan nonlinier

dengan menginstruksikan data berdasarkan teori statistika tradisional (Mohan et al., 2020). Metode ini bekerja dengan membuat suatu pembatas keputusan yang memisahkan atau mengklasifikasikan semua vektor pelatihan ke dalam subkelas (Mohan et al., 2020).

Support Vector Machine bekerja dengan cara menemukan *hyperplane* terbaik yang dapat memisahkan dua kelas dengan mencari jarak terjauh di antara keduanya (Kurniadi et al., 2021). SVM menemukan *hyperplane* ini menggunakan vektor pendukung dan margin (Mohan et al., 2020). *Hyperplane* berfungsi sebagai pemisah antara kelas-kelas tersebut, di mana dalam kasus fungsi 2 dimensi, disebut sebagai *line whereas*, sedangkan dalam kasus fungsi 3 dimensi, disebut sebagai *plane similarly* (Kurniadi et al., 2021). Untuk klasifikasi dalam ruang dimensi yang lebih tinggi dari 3 dimensi, *hyperplane* digunakan sebagai pemisah antara kelas-kelas tersebut (Kurniadi et al., 2021). Ilustrasi pada Gambar 1 berikut menggambarkan metode SVM.



Gambar 1. Ilustrasi metode SVM (Kurniadi et al., 2021)

Kernel digunakan untuk mengubah ruang nonlinear menjadi ruang *linear* (Mohan et al., 2020). Kernel merupakan sebuah komponen yang dapat menjalankan fungsi pemetaan dari input berdimensi rendah ke dalam ruang dimensi yang lebih tinggi (Mohan et al., 2020). Kernel berfungsi untuk menerima data dan mengkonversinya ke format yang sesuai (Almaiah et al., 2022). Menurut (Mohan et al., 2020) dan (Almaiah et al., 2022) beberapa

fungsi kernel yang digunakan dalam SVM meliputi linear, polynomial, RBF, dan sigmoid.

a. *Linear kernel*

Linear kernel adalah jenis kernel yang lebih tradisional dan lebih mudah digunakan. Fungsi ini disajikan pada persamaan (1).

$$k(x, y) = x^T y \quad (1)$$

Keterangan:

x^T dan y = vektor input.

Penjumlahan perkalian dalam $\{x, y\}$ dan konstanta lainnya. Fungsinya adalah untuk memungkinkan pemisahan data secara linier, terutama cocok untuk *dataset* yang relatif kecil dengan waktu pelatihan yang singkat.

b. *Polynomial Kernel*

Polynomial kernel didefinisikan sebagai kernel non-stasioner, dan biasanya cocok pada set data pelatihan yang telah dinormalisasi. Parameter yang dapat disesuaikan dikenal sebagai kemiringan alfa, konstanta c , dan derajat polinomial. Fungsi polinomial kernel disajikan pada persamaan (2) berikut.

$$(x, y) = (x^T y + 1)^d \quad (2)$$

Keterangan:

x^T dan y = vektor input.

d = derajat polinomial.

Polinomial kernel digunakan ketika terdapat sedikit sampel atau kurangnya variasi. Dalam penggunaannya, kernel polinomial menawarkan fitur global yang berkualitas baik serta cakupan area konvergensi yang luas.

c. *Radial basis function*

Kernel Gaussian digunakan untuk menganalisis dan memprediksi hasil dalam laporan berdasarkan sekumpulan data. Kernel Gaussian adalah contoh dari kernel berbasis radial kernel. Fungsi Kernel Gaussian disajikan pada persamaan (3) berikut.

$$(x, y) = \exp\left(-\frac{\|x - y\|^2}{2\sigma^2}\right) \quad (3)$$

Keterangan:

x dan y = vektor input.

$\|x-y\|$ = jarak euclidean kuadrat antara dua vektor.

Sigma = parameter bandwidth yang mengontrol kelancaran batas keputusan.

Kernel RBF memiliki kemampuan untuk menangani fitur lokal dengan baik dan memiliki cakupan area konvergensi yang luas, cocok untuk digunakan pada *dataset* besar maupun kecil.

d. *Hyperbolic Tangent Kernel (Sigmoid Kernel)*

Ketika ruang fitur memiliki dimensi yang rendah *Hyperbolic Tangent Kernel* memiliki kecenderungan untuk konvergen dengan baik, baik untuk sampel kecil maupun besar. Jenis kernel ini cukup terkenal dalam mesin vektor, dan perkembangannya dipengaruhi oleh jaringan syaraf. Fungsi kernel ini dapat direpresentasikan pada persamaan (4) berikut.

$$(x, y) = \tanh(ax^T y + c) \quad (4)$$

Keterangan:

x^T dan y = vektor input.

a = parameter skala yang mengontrol kemiringan fungsi.

c = konstanta yang menggeser fungsi.

\tanh = fungsi tangen hiperbolik.

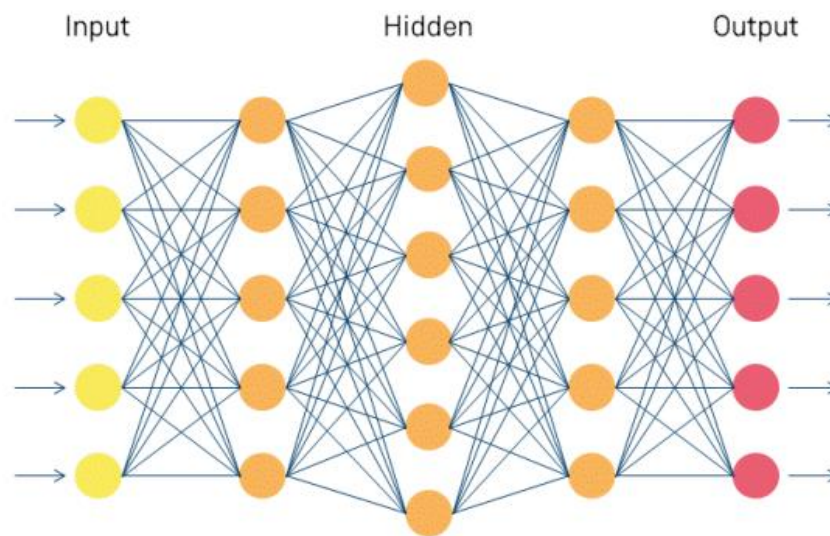
Untuk sampel besar maupun kecil, *hyperbolic tangent kernel* memiliki kecenderungan konvergensi yang baik saat dimensi ruang fitur rendah. Namun, dalam kasus dimensi ruang yang tinggi, mungkin terdapat peluang yang berbeda dalam mencapai konvergensi.

2.4 Deep Learning

Metode *deep learning* adalah sebuah metode yang berkembang dari *machine learning*. *Deep learning* merupakan pendekatan pemodelan berbasis komputer yang terdiri dari banyak lapisan pemrosesan yang digunakan untuk memahami representasi data pada beberapa tingkatan abstraksi (Mishra et al., 2021). *Deep learning* adalah salah satu cabang dari *machine learning* yang memiliki kinerja jauh lebih baik untuk menangani data tidak terstruktur (Mathew et al., 2021). Saat ini, teknologi *deep learning* telah mengungguli kemampuan dari *machine learning* (Mathew et al., 2021). *Deep learning* memungkinkan model komutasi untuk mempelajari fitur secara bertahap dari suatu data dengan berbagai tingkatan yang berbeda (Mathew et al., 2021).

Prinsip kerja dari *deep learning* adalah dengan mengekstraksi fitur dari data mentah dengan menggunakan beberapa lapisan yang digunakan untuk mengidentifikasi berbagai aspek yang relevan dengan data masukan (Mishra et al., 2021). *Deep learning* sangat baik dalam menemukan struktur yang rumit dalam data berdimensi tinggi, memecahkan masalah seperti pengenalan gambar, pengenalan suara, dan memprediksi aktivitas molekul obat (LeCun et al., 2015).

Deep learning mengidentifikasi pola-pola kompleks dalam *dataset* yang besar menggunakan algoritme *backpropagation* untuk menentukan bagaimana sebuah mesin harus menyesuaikan parameter internalnya agar dapat menghasilkan representasi di setiap lapisan berdasarkan representasi dari lapisan sebelumnya (LeCun et al., 2015). Jaringan konvolusi telah menghasilkan terobosan dalam pemrosesan gambar, video, ucapan, dan audio, sedangkan jaringan *recurrent* telah memberikan kemajuan untuk data berurutan seperti teks dan ucapan (LeCun et al., 2015). Gambar 2 berikut merupakan gambaran contoh jaringan *deep learning* yang terdiri dari tiga lapisan serta memiliki *hidden layer*.



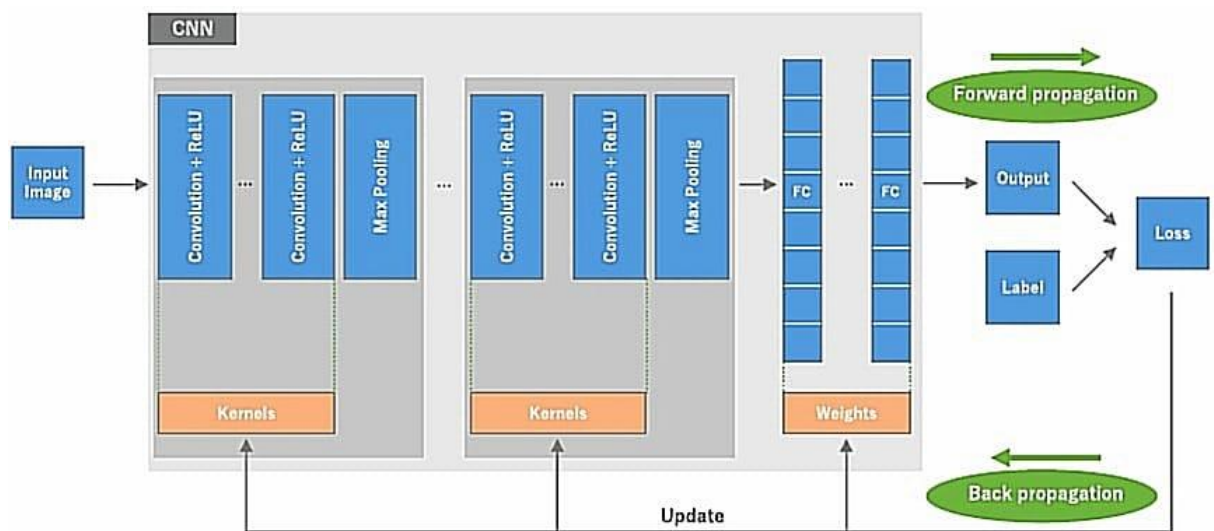
Gambar 2. Contoh jaringan *deep learning* (Nahr et al., 2021).

2.5 Convolutional Neural Network

Convolutional neural network adalah salah satu jenis jaringan syaraf tiruan yang diterapkan dalam *deep learning* yang sebagian besarnya digunakan untuk menganalisis gambar atau data visual (Agarwal & Singh, 2022). *Convolutional neuron* tersusun dari bobot dan bias yang dapat dipelajari dan disesuaikan untuk mencapai hasil yang diinginkan (Agarwal & Singh, 2022). Ribuan dan jutaan neuron ini ketika disusun dalam urutan tertentu dengan *neuron* lain menghasilkan jaringan syaraf tiruan yaitu CNN (Agarwal & Singh, 2022). Pada umumnya CNN digunakan untuk mengklasifikasikan dan mengurutkan data serta gambar dan mengelompokkan yang tampak serupa kemudian melakukan pengenalan objek (Agarwal & Singh, 2022). *Convolutional neural network* (CNN) adalah jenis pembelajaran mendalam yang secara otomatis dapat mempelajari sekumpulan detektor fitur, melalui serangkaian lapisan yang membentuk representasi model yang mendalam, dari kumpulan data berlabel (Yamashita et al., 2018). Metode CNN telah memperoleh banyak keberhasilan karena pada CNN ketika data dan kompleksitas model

meningkat, jaringan saraf akan tumbuh dan beradaptasi sesuai dengan data tersebut (Nahr et al., 2021).

CNN telah mendominasi dalam berbagai kelas komputer vision, menarik minat di berbagai domain, termasuk radiologi (Yamashita et al., 2018). CNN dirancang secara otomatis dan adaptif mempelajari hirarki spasial fitur dari pola tingkat rendah hingga tinggi (Yamashita et al., 2018). CNN terdiri dari tiga lapisan atau menggunakan beberapa blok bangunan, seperti lapisan konvolusi (*convolutional layer*), lapisan penyatuan (*pooling layer*), dan lapisan yang terhubung sepenuhnya (*fully connected layers*) (Yamashita et al., 2018). Dua lapisan pertama yaitu lapisan *convolutional* dan *pooling layer* berfungsi untuk ekstraksi fitur, sedangkan lapisan *fully connected layer* berfungsi untuk memetakan fitur yang diekstraksi menjadi hasil akhir seperti klasifikasi (Yamashita et al., 2018). Gambar 3 berikut merupakan gambaran umum arsitektur CNN dan proses *training* dalam CNN.

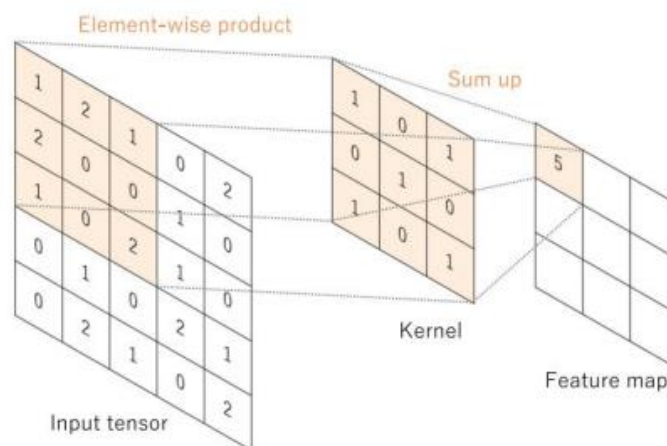


Gambar 3. *Convolutional Neural Network* (Yamashita et al., 2018).

2.5.1 Layer CNN

a. *Convolutional Layer*

Convolutional layer merupakan komponen dasar dari arsitektur CNN dengan mengekstraksi fitur, yang biasanya terdiri dari kombinasi operasi linier dan nonlinier, yaitu operasi konvolusi dan fungsi aktivasi (Yamashita et al., 2018). Konvolusi termasuk ke dalam jenis operasi linear khusus yang digunakan untuk ekstraksi fitur di mana sejumlah kecil kota yang berisi angka disebut kernel diterapkan pada data inputan yang merupakan larik angka yang disebut tensor (Yamashita et al., 2018). Proses ini membantu dalam menemukan pola maupun fitur khusus dalam data. Dengan kata lain konvolusi dapat disebut cara untuk melihat bagian bagian kecil untuk memahami gambaran besar dari data yang ada atau mengekstrak informasi penting dari data (Yamashita et al., 2018).

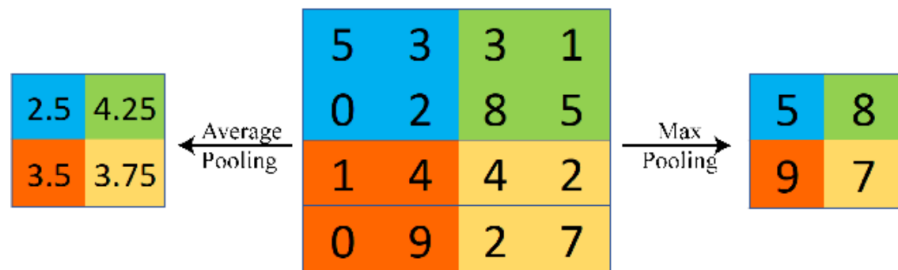


Gambar 4. *Convolutional Layer* (Yamashita et al., 2018).

b. *Pooling Layer*

Pooling layer atau lapisan penyatuan menyediakan operasi *downsampling* yang mengurangi dimensi dalam bidang peta bidang peta fitur untuk memperkenalkan invarian terjemahan ke pergeseran kecil (Yamashita et al., 2018). *Pooling* adalah salah satu lapisan penting dalam jaringan saraf tiruan yang bertanggung jawab untuk pengurangan dimensi atau ukuran spasial sari matriks fitur sehingga membantu mengurangi daya komputasi yang diperlukan untuk melatih *neural network* (Agarwal & Singh, 2022). Ada dua tipe

utama dari *pooling layers* yang dapat digabung dalam *deep neural network* di antaranya *Max Pooling* dan *Average Pooling*.



Gambar 5. *Average Pooling vs Max Pooling* (Agarwal & Singh, 2022).

Max Pooling, seperti namanya, memberikan nilai maksimum untuk area gambar yang dicakup oleh kernel, sedangkan *Average Pooling* mengembalikan rata-rata nilai untuk area gambar yang dicakup oleh kernel (Agarwal & Singh, 2022). Bentuk *pooling operation* paling populer adalah *max-pooling* yang mengekstrak fitur dari *input feature maps*, mengeluarkan nilai maksimum di setiap *patch* dan membuang *patch* lainnya (Chen et al., 2018). *Max pooling* membantu dalam menghilangkan *noise* atau fitur yang mengganggu dengan pengurangan dimensi, sedangkan *Average Pooling* hanya membantu pengurangan dimensi (Agarwal & Singh, 2022).

Dengan demikian pada penelitian ini saya menggunakan *Max Pooling*. *Pooling layer* sangat penting dalam membantu pengambilan informasi penting dari lautan data untuk membantu model mempelajari fitur-fitur pada *dataset* (Agarwal & Singh, 2022). Penghapusan *noise* dan penyediaan fitur yang berguna dapat membantu mencegah *overfitting* dan mempercepat proses komputasi (Agarwal & Singh, 2022).

c. *Fully Connected Layer*

Peta fitur keluaran dari konvolusi akhir atau *pooling layer* biasanya diratakan, yaitu, diubah menjadi susunan angka (atau vektor) satu

dimensi (1D), dan terhubung ke satu atau lebih *fully connected layer*, yang juga dikenal sebagai *dense layer*, di mana setiap input terhubung ke setiap output dengan bobot yang dapat dipelajari (Yamashita et al., 2018). Setelah fitur-fitur yang diekstraksi oleh lapisan konvolusi dan diambil sampelnya oleh *pooling layer* dibuat, fitur-fitur tersebut diproses oleh subset dari *fully connected layer* menuju *output* akhir jaringan, seperti menghitung probabilitas untuk masing-masing kelas dalam tugas klasifikasi (Yamashita et al., 2018). Lapisan akhir dari *fully connected layer* biasanya memiliki jumlah *node output* yang sama dengan jumlah kelas (Yamashita et al., 2018). Setiap *fully connected layer* diikuti oleh fungsi non-linier, seperti ReLU (Yamashita et al., 2018).

2.5.2 *Rectified Linear Unit (ReLU)*

Fungsi Aktivasi juga disebut sebagai lapisan pemetaan *non-linear* yang biasa digunakan untuk meningkatkan kemampuan jaringan klasifikasi (Chen et al., 2018). Fungsi aktivasi secara khusus digunakan dalam jaringan syaraf tiruan untuk mengubah sinyal input menjadi sinyal *output* yang kemudian menjadi lapisan *input* untuk lapisan berikutnya (Sharma et al., 2020).

Rectified linear unit (ReLU) merupakan salah satu fungsi aktivasi yang sering digunakan pada *Convolutional Neural Network* (Chen et al., 2018). Keuntungan menggunakan fungsi ReLU adalah semua *neuron* tidak diaktifkan pada waktu yang sama (Chen et al., 2018). Hal ini menyiratkan bahwa sebuah neuron akan dinonaktifkan hanya ketika output dari transformasi linier adalah nol. Fungsi ReLU adalah *piecewise function* yang didefinisikan pada persamaan (5) berikut:

$$\text{Rectifier}(x) = \{\max(0, x)\} = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (5)$$

ReLU lebih efisien dibandingkan dengan fungsi lainnya karena semua neuron tidak diaktifkan pada waktu yang sama, melainkan sejumlah neuron diaktifkan pada satu waktu (Sharma et al., 2020). Dalam

beberapa kasus, nilai gradien adalah nol, yang menyebabkan bobot dan bias tidak diperbarui selama langkah *backpropagation* dalam pelatihan jaringan syaraf (Sharma et al., 2020).

2.5.3 *Batch Normalization*

Batch Normalization adalah proses normalisasi efek pada setiap layer untuk setiap batch pelatihan. Teknik ini berfungsi untuk mengurangi kebutuhan daya komputasi dan jumlah epoch pelatihan dalam model *deep learning* (Agarwal & Singh, 2022). Selama pelatihan jaringan neural, input pada layer dapat mengalami perubahan karena pembaruan bobot di setiap mini batch. Kondisi ini bisa menimbulkan *internal covariate shift*, yaitu perubahan distribusi *input* pada *layer* untuk tiap *mini batch* (Agarwal & Singh, 2022).

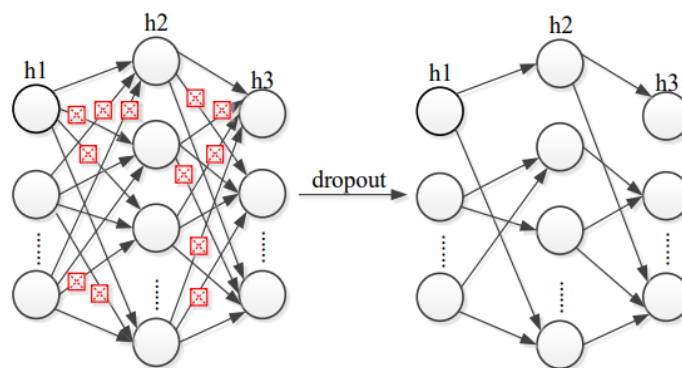
2.5.4 *Flatten*

Flatten merupakan metode yang digunakan untuk mengubah data menjadi larik satu dimensi sebagai masukan untuk lapisan berikutnya (Agarwal & Singh, 2022). *Output* berupa vektor atau matriks fitur diproses menjadi vektor tunggal yang dapat dimanfaatkan sebagai *input* ke lapisan berikutnya untuk analisis lebih lanjut (Agarwal & Singh, 2022).

2.5.5 *Dropout*

Dropout layer memiliki fungsi untuk mengurangi tingkat *overfitting* dalam jaringan yang besar (Chen et al., 2018). *Overfitting* merupakan masalah yang cukup besar dalam *deep learning* sehingga berbagai metode telah dirancang oleh peneliti (Chen et al., 2018). Secara sederhana, sebuah *dropout layer* akan mengabaikan sejumlah neuron secara acak dalam proses *training* untuk mencegah *overfitting* dan bias pada *output* yang dihasilkan model (Agarwal & Singh, 2022). *Dropout layer* dapat membuat struktur jaringan menjadi lebih bersih dan teratur (Chen et al., 2018). Korelasi antar *neuron* berkurang setelah konvolusi

dan jaringan bisa mendapatkan parameter yang lebih baik dalam proses memperbaharui bobot (Chen et al., 2018).



Gambar 6. *Dropout Layer* (Chen et al., 2018).

2.6 Image Augmentation

Augmentasi data adalah teknik yang dapat digunakan untuk memperbesar *dataset* untuk meningkatkan keragaman dan memperluas data sehingga dapat menghasilkan model yang lebih baik (Agarwal & Singh, 2022). Teknik augmentasi data dapat digunakan untuk meningkatkan ukuran dan kualitas *dataset* pelatihan untuk model *deep learning*, sehingga dapat meningkatkan performa model dengan data terbatas seperti analisis citra medis (Shorten & Khoshgoftaar, 2019). Augmentasi data bertujuan untuk menghasilkan data tambahan yang digunakan untuk melatih model dan telah terbukti meningkatkan kinerja ketika divalidasi pada *dataset* yang tidak terlihat (Chlap et al., 2021).

Teknik augmentasi dapat dibedakan menjadi dua jenis yaitu *basic augmentation techniques* dan *deformable augmentation techniques* (Chlap et al., 2021). Teknik augmentasi dasar meliputi penerapan transformasi pada gambar yang mengubah posisi titik-titik gambar atau memanipulasi nilai intensitas gambar, dengan tujuan menghasilkan gambar yang diperbesar (Chlap et al., 2021). Sementara itu, teknik augmentasi yang bersifat *deformable* dapat diterapkan ketika teknik augmentasi dasar tidak

memberikan variasi yang memadai untuk memastikan generalisasi yang optimal pada model berikutnya (Chlap et al., 2021).

Pada penelitian ini hanya berfokus pada peningkatan variasi sederhana dan pemrosesan gambar yang cepat sehingga yang akan digunakan adalah *basic augmentation techniques*. Proses augmentasi dilakukan pada satu gambar dari *dataset* yang ada, dan setelah mengalami modifikasi, gambar tersebut dimasukkan kembali ke dalam *dataset* dengan tujuan meningkatkan ukuran keseluruhan dari *dataset* tersebut (Chlap et al., 2021). Meskipun teknik-teknik ini simpel dalam penerapannya, namun memiliki potensi efektivitas yang tinggi dalam meningkatkan kinerja model yang sedang dilatih (Chlap et al., 2021). Menurut (Chlap et al., 2021) *basic augmentation techniques* di antaranya sebagai berikut.

a. *Geometric transformations*

Transformasi geometri adalah bentuk paling umum dalam teknik augmentasi di antaranya ada *scaling*, *translating*, *rotating*, *reflecting*, *zooming*, *shearing* atau terkadang transformasi perspektif termasuk kemiringan.

b. *Cropping*

Cropping merupakan teknik di mana potongan-potongan acak dipilih dari sebuah gambar yang sudah ada. Potongan-potongan acak ini kemudian ditambahkan kembali ke dalam *dataset* untuk meningkatkan ukuran, khususnya digunakan ketika terdapat ketidakseimbangan kelas.

c. *Occlusion*

Occlusion melibatkan penghapusan potongan-potongan gambar untuk menghasilkan gambar yang telah di augmentasi.

d. *Intensity operations*

Intensity operations memanipulasi nilai piksel/*voxel* dalam gambar. Dilakukan dengan memodifikasi kecerahan atau kontras gambar dengan metode yang umum digunakan adalah koreksi gamma, kontras *linear*, dan ekualisasi histogram.

e. *Noise injection*

Noise injection adalah teknik augmentasi data yang umum digunakan untuk mensimulasikan gambar dengan tambahan *noise*. Penyisipan *noise gaussian* adalah yang paling umum, di mana intensitas gambar dimodifikasi dengan secara acak mengambil sampel dari distribusi Gaussian. Selain itu metode yang digunakan di antaranya *uniform noise* yang memodifikasi nilai dengan secara acak mengambil sampel dari *uniform distribution* atau *noise salt* dan *pepper*, dimana piksel secara acak diatur menjadi hitam atau putih

f. *Filtering*

Pemfilteran suatu gambar dilakukan dengan menggunakan konvolusi. Untuk mencapai ini, sebuah kernel konvolusi digerakkan melintasi gambar untuk memodifikasi intensitas di setiap piksel berdasarkan nilai piksel sekitarnya. Dengan menggunakan metode ini, sebuah gambar dapat diperjelas, diburamkan, atau diratakan sehingga menghasilkan gambar yang telah di augmentasi.

g. *Combination*

Combination adalah metode augmentasi data yang menghasilkan gambar baru dengan menggabungkan dua atau lebih gambar asli. Dua penelitian menggunakan teknik *mix-up*, yang bekerja secara acak memilih dua gambar dari kumpulan data dan menggabungkan intensitas *voxel* yang sesuai dari kedua gambar tersebut.

Pada penelitian ini teknik augmentasi yang akan digunakan adalah *geometric transformations* dengan menggunakan 3 teknik yaitu *rotating* (menggeser gambar sejauh apa), *flip_horizontal* (membalik gambar secara horizontal) dan *zooming* (memperbesar gambar).

2.7 Confusion Matrix

Confusion matrix adalah suatu metode yang digunakan dalam pembelajaran mesin untuk mengevaluasi atau memvisualisasikan bagaimana kinerja

model dalam konteks klasifikasi (Caelen, 2017). *Confusion matrix* merupakan matriks bujur sangkar di mana baris mewakili kelas aktual dari instance dan kolom mewakili kelas prediksi mereka (Caelen, 2017). Dalam menangani tugas klasifikasi biner, maka matriks konfusi adalah matriks 2×2 yang melaporkan jumlah *true positives*, *true negatives*, *false positives*, dan *false negatives* (Caelen, 2017). Berdasarkan penelitian sebelumnya (Agarwal & Singh, 2022) akurasi model, *presisi*, *recall*, dan skor F1 adalah parameter *confusion matrix* yang digunakan untuk mengukur kinerja model.

a. *Accuracy*

Akurasi mencerminkan sejauh mana sistem mampu mengklasifikasikan data dengan tepat. Ini dihitung dengan membandingkan jumlah data yang diklasifikasikan dengan benar dengan keseluruhan jumlah data (Agustina et al., 2022). Akurasi dirumuskan pada Persamaan (6) berikut.

$$Accuracy = \frac{True\ Positive + True\ Negative}{True\ Positive + False\ Positive + False\ Negative + True\ Negative} \quad (6)$$

b. *Precision*

Presisi mengukur rasio antara jumlah data yang benar terklasifikasi dalam suatu kategori dengan total keseluruhan data yang terklasifikasi dalam kategori tersebut. (Agustina et al., 2022). *Precision* dirumuskan pada Persamaan (7) sebagai berikut.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (7)$$

c. *Recall*

Recall, dilakukan untuk mengevaluasi perbandingan antara jumlah data yang benar terklasifikasi oleh sistem dalam suatu kategori dengan total gabungan data yang benar dan salah terklasifikasi dalam kategori tersebut (Agustina et al., 2022). *Recall* dirumuskan pada Persamaan (8) berikut.

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (8)$$

d. *F1-Score*

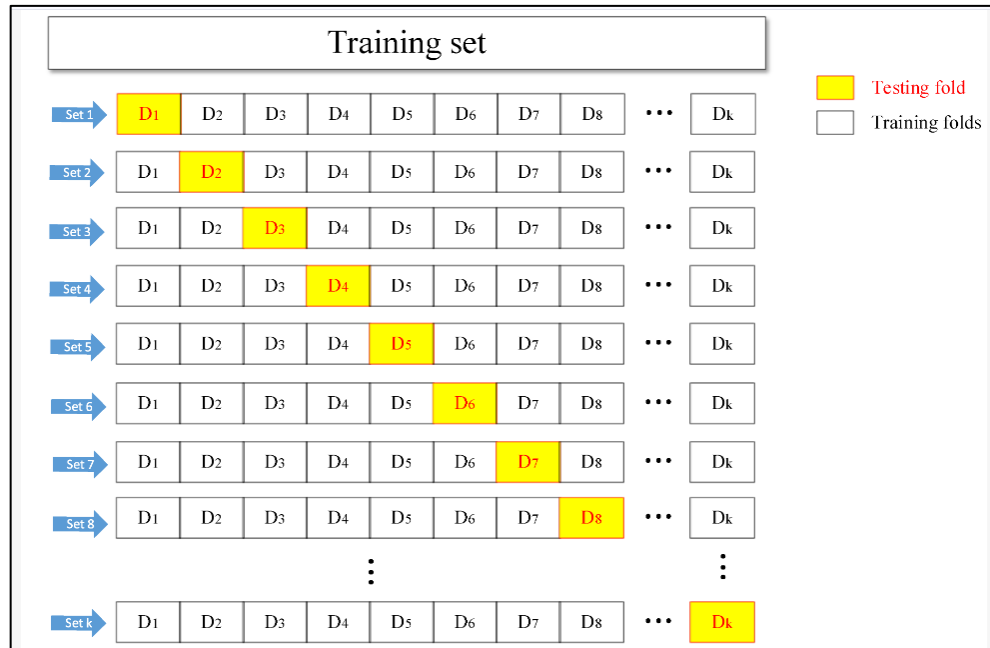
F1-Score adalah kombinasi dari presisi dan *recall*. *F1-Score* mencerminkan keseimbangan antara ketepatan (*precision*) dan (*recall*) dalam mengukur kinerja suatu sistem klasifikasi (Agustina et al., 2022).

F1-Score dirumuskan pada Persamaan (9) berikut.

$$F1-Score = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (9)$$

2.8 *K-Fold Cross Validation*

Salah satu aspek penting dalam penelitian adalah cara membagi data menjadi set pelatihan dan pengujian untuk mengukur kinerja pengklasifikasi. Pembagian data dapat dilakukan dengan menggunakan metode validasi silang. Teknik validasi silang yang paling umum digunakan adalah *k-fold cross-validation*, di mana *dataset* dibagi secara acak menjadi *k* subset, setiap subset secara bergantian digunakan sebagai data pengujian, sementara *k-1* subset lainnya digunakan untuk pelatihan (White & Power, 2023). Proses ini diulangi sebanyak *k* kali, sehingga setiap subset berperan sebagai data pengujian sekali (White & Power, 2023). Kinerja keseluruhan pengklasifikasi kemudian diukur dengan menghitung rata-rata dari *k* akurasi klasifikasi yang dihasilkan pada setiap iterasi. Variasi akurasi dapat terjadi akibat pembagian pelatihan/pengujian yang berbeda, metode *k-fold cross validation* memberikan estimasi kinerja pengklasifikasi yang lebih akurat dan dapat digeneralisasi dibandingkan dengan hanya menggunakan satu pembagian data (White & Power, 2023). Berikut ini gambaran proses pembagian data menggunakan *k-fold cross validation*.



Gambar 7. Proses *k-fold cross validation* (Lyu et al., 2022).

2.9 Kanker Kulit

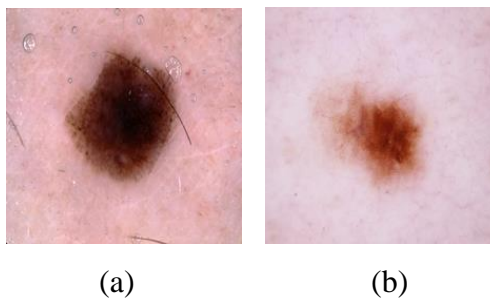
Kulit adalah organ tubuh terbesar yang berfungsi melindungi seluruh organ dalam tubuh dari dunia luar yang berbahaya. Kulit membantu mengatur suhu serta melindungi tubuh dari infeksi (Agarwal & Singh, 2022). Kulit terdiri dari tiga lapisan yaitu Epidermis, Dermis, dan Hipodermis (Agarwal & Singh, 2022). Kulit merupakan organ yang sangat penting untuk tubuh manusia, namun kulit juga dapat terserang penyakit berbahaya yang dapat mengganggu fungsinya. Salah satu penyakit kulit berbahaya yang dapat mengancam nyawa adalah kanker.

Kanker kulit merupakan jenis kanker yang paling sering terjadi di seluruh dunia, dan angka kasusnya terus meningkat (Craythorne & Al-Niami, 2017). Hal ini terjadi karena ini pertumbuhan sel yang cepat (Waheed et al., 2017). Kanker kulit adalah bentuk kanker yang umumnya dikaitkan dengan paparan sinar matahari yang berlebihan atau bentuk partikel dengan energi tinggi yang menghancurkan susunan genetik sel (Agarwal & Singh,

2022). Risiko terkena kanker kulit muncul dari kombinasi faktor genetik dan lingkungan, penyebab paling umum adalah paparan sinar ultraviolet (UV) dalam waktu lama (Craythorne & Al-Niami, 2017). Kanker kulit menjadi sebagai salah satu jenis kanker yang paling berbahaya dan mengalami peningkatan tingkat kematian drastis karena kurangnya pengetahuan tentang gejala dan pencegahannya (Monika et al., 2020).

Menurut (Agarwal & Singh, 2022) kanker dapat dideteksi apabila massa sel kulit yang sehat berubah sifat dan mulai tumbuh tak terkendali membentuk massa yang disebut tumor. Pola pertumbuhan sel yang tidak teratur ini dapat dikategorikan sebagai jinak atau ganas (Monika et al., 2020). Tumor ini dapat berupa pertumbuhan jinak yang berarti tidak membahayakan kehidupan dan tidak tumbuh menyebar ke daerah lain, meskipun bisa saja menjadi ganas yang mengganggu dan dapat tumbuh menyebar ke bagian tubuh lainnya (Agarwal & Singh, 2022).

Kanker kulit secara luas dibagi menjadi kanker yang berasal dari *melanosit* (melanoma) dan dari sel yang berasal dari epidermis (kanker kulit non-melanoma). Kedua kelompok ini merupakan mayoritas kanker kulit (95%), sedangkan tumor kulit lainnya hanya merupakan persentase yang sangat kecil (Craythorne & Al-Niami, 2017). Dari dua jenis kanker tersebut dapat dibagi lagi menjadi 3 jenis yaitu tiga tipe utama yaitu *Basal Cell Carcinoma* (BCC), *Squamous* (SCC) Melanoma (Waheed et al., 2017). BCC terjadi ketika lesi tumbuh dengan cepat di sel kulit disebut sebagai sel basal, SCC mempengaruhi sel kulit *skuamosa* dan melanoma terjadi pada sel kulit *melanosit* (Waheed et al., 2017). Kanker kulit melanoma dapat diklasifikasikan secara luas ke dalam dua kategori utama: melanoma (ganas) dan non-melanoma (Jinak) (Agarwal & Singh, 2022). Kanker kulit jinak dan ganas digambarkan pada Gambar 8 berikut.



Gambar 8. (a) Kanker kulit ganas, (b) Kanker kulit jinak
(www.kaggle.com).

Di antara berbagai jenis kanker kulit, *melanoma malignant* adalah jenis kanker kulit yang paling agresif dan mematikan (Waheed et al., 2017). Beberapa gejala awal yang terkait dengan kanker ini adalah perubahan warna atau peradangan pada kulit, gatal dan pendarahan pada bercak kulit, tahi lalat atau merah, benjolan seperti lilin pada kulit akibat sel kanker (Agarwal & Singh, 2022). Deteksi melanoma kanker pada tahap awal dapat meningkatkan kemungkinan bertahan hidup (Agarwal & Singh, 2022). Pasien melanoma dapat disembuhkan jika penyakit ini terdeteksi sejak dini (Waheed et al., 2017).

Biasanya, melanoma muncul sebagai lesi berpigmen baru atau mengalami perubahan, terutama pada area kulit yang sering terpapar sinar matahari pada individu dengan warna kulit putih (Craythorne & Al-Niami, 2017). Melanoma memiliki beberapa subtype yang berbeda (Craythorne & Al-Niami, 2017). Jika dapat dideteksi dan diobati pada tahap awal, prospek penyembuhannya sangat baik. Namun, prognosisnya menjadi lebih buruk dengan cepat jika lesi tersebut semakin tebal (Craythorne & Al 2017).

Tumor jinak adalah jenis kanker yang umumnya dianggap sebagai tahi lalat, yang tidak berbahaya, sedangkan tumor ganas dianggap sebagai kanker yang mengancam nyawa karena dapat merusak jaringan tubuh lainnya (Monika et al., 2020). Deteksi kanker kulit pada stadium dini adalah cara yang terbaik untuk mengurangi dampak penyakit tersebut (Monika et al., 2020).

III. METODE PENELITIAN

3.1 Tempat dan Waktu

3.1.1 Tempat penelitian

Penelitian ini dilaksanakan di Laboratorium Komputasi Dasar Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung, yang berlokasi di Jalan Prof. Dr. Ir. Sumantri Brojonegoro No. 1, Gedong Meneng, Kecamatan Rajabasa, Kota Bandar Lampung, Lampung 35141.

3.1.2 Waktu Penelitian

Waktu pelaksanaan penelitian dimulai dari bulan Januari 2024 dengan rencana jadwal penelitian yang dapat dilihat pada Tabel 2.

Tabel 2. Jadwal penelitian

Kegiatan	2024																															
	Januari				Februari				Maret				April				Mei				Juni				Juli				Agustus			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Studi literatur	■	■																														
Pengumpulan data			■	■																												
penyusunan BAB I-III					■	■	■	■	■	■	■	■																				
<i>Preprocessing</i>									■	■	■	■																				
Pembagian data													■	■																		
Pemodelan menggunakan CNN																	■	■	■	■	■	■										
Ekstraksi Fitur SVM																					■	■										
Pemodelan menggunakan SVM																					■	■	■	■								
Evaluasi Kinerja																									■	■	■	■	■	■	■	■
Penulisan hasil BAB IV-V																									■	■	■	■	■	■	■	■

3.2 Alat dan Bahan

Alat dan bahan yang digunakan dalam penelitian ini berupa *hardware* dan *software* sebagai berikut.

3.2.1 *Hardware*

Hardware yang digunakan dalam penelitian ini berupa laptop *Asus VivoBook Max* dengan spesifikasi sebagai berikut.

- a. *Processor AMD Dual Core A9-942* kecepatan 3.7 Hz.
- b. RAM 4 GB.
- c. HDD 1T.

3.2.2 *Software*

Software yang digunakan dalam penelitian ini antara lain sebagai berikut.

- a. Sistem operasi Windows 10.
- b. Text Editor Google Colaboratory, untuk menjalankan kode secara cloud.
- c. Bahasa pemrograman menggunakan Python.
- d. Penyimpanan data menggunakan Google Drive.

3.2.3 *Library*

Library yang digunakan dalam penelitian ini diantaranya sebagai berikut.

a. Scikit-learn

scikit-learn adalah *library* dalam Python yang mendukung berbagai algoritma pembelajaran mesin seperti klasifikasi, regresi, dan klustering. Selain itu, *library* ini juga menyediakan alat untuk memilih dan mengevaluasi model (Sundaram et al., 2023).

b. Numpy

NumPy adalah salah satu *library* paling populer untuk komputasi ilmiah. *Library* numpy menyediakan operasi matematika pada array

dan matriks tersedia pula fungsi di *library* Open cv (Sundaram et al., 2023).

c. Scikit-image

Scikit-image atau skimage dirancang untuk mempermudah proses pengolahan gambar pada python, terutama bagi yang terbiasa dengan *tools* python. Salah satu cara untuk mencapai hal ini adalah dengan menggunakan array NumPy standar sebagai representasi dasar dari gambar yang memungkinkan akses langsung ke data piksel (van der Walt et al., 2014).

d. Tensorflow

TensorFlow dalam Python menyediakan dukungan untuk berbagai algoritma pembelajaran mesin (*machine learning*) dan pembelajaran mendalam (*deep learning*). Program ini mencakup berbagai *tools*, termasuk untuk mengevaluasi dan menerapkan model, serta *tools* untuk membuat dan melatih jaringan saraf (Sundaram et al., 2023).

e. Keras

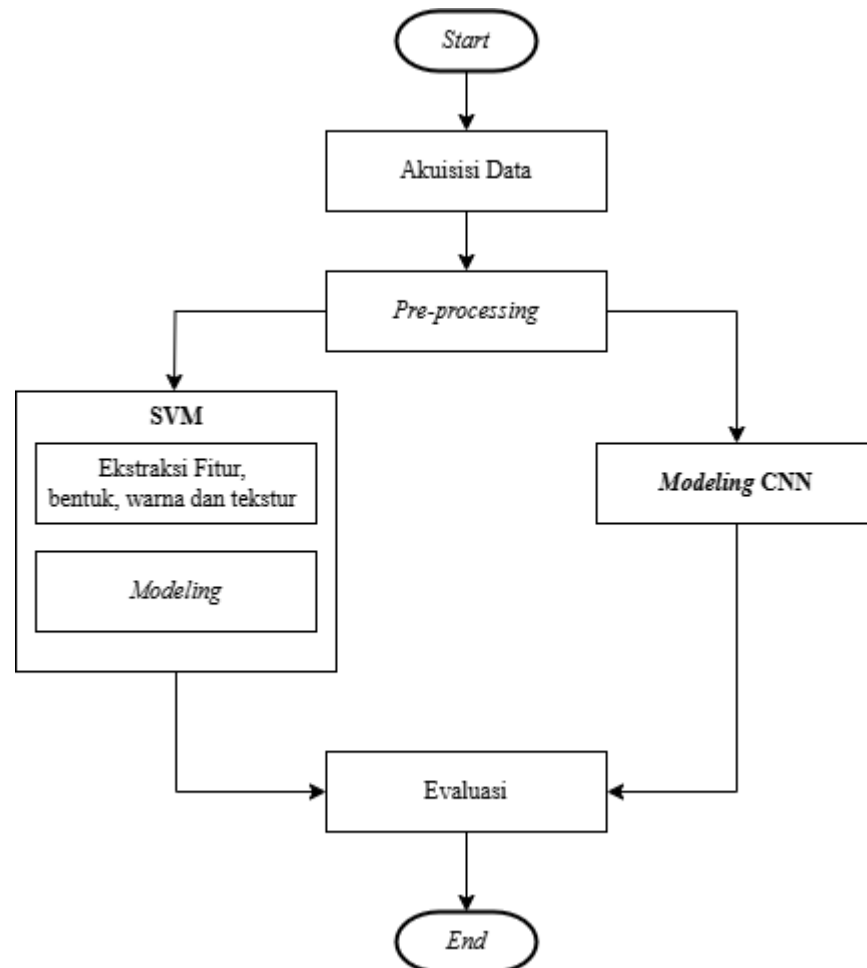
Keras adalah sebuah pustaka jaringan saraf tingkat tinggi berbasis Python yang digunakan di platform TensorFlow, Theano, dan CNTK untuk membangun jaringan saraf yang kuat. Keras memiliki antarmuka yang mudah digunakan, yang menjadikan proses pembuatan model pembelajaran mendalam lebih sederhana (Sundaram et al., 2023)..

f. Matplotlib.

Perpustakaan Matplotlib dalam Python digunakan untuk membuat visualisasi statis, animasi, dan visualisasi interaktif menggunakan Python. Perpustakaan ini menyediakan berbagai jenis grafik yang dapat disesuaikan dan tersedia untuk banyak format (Sundaram et al., 2023).

3.3 Tahap Penelitian

Tahap penelitian yang dilakukan dalam penelitian ini digambarkan pada Gambar 9 sebagai berikut.



Gambar 9. Alur kerja penelitian.

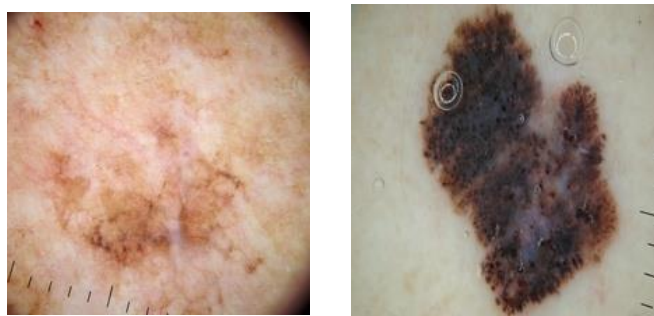
3.3.1 Akuisisi Data

Data yang digunakan diperoleh dari Kaggle *dataset* <https://www.kaggle.com/fanconic/skin-cancer-malignant-vs-benign> dengan judul *dataset* “Skin Cancer: Malignant vs. Benign”. *Dataset* berisi dua kelas citra kanker kulit yaitu ganas (*malignant*) dan jinak (*benign*). Data tersebut memiliki jumlah total data sebanyak 3297.

Menurut sumber, data tersebut diperoleh dari arsip *International Skin Imaging Collaboration (ISIC)*. Rincian data yang digunakan dalam penelitian ini adalah sebanyak 1800 jenis citra jinak (*benign skin moles*) dan 1497 jenis data ganas (*malignant skin moles*). Sampel citra *dataset* tersebut dapat dilihat pada Gambar 10 dan 11 berikut.



Gambar 10. Kanker kulit jinak (www.kaggle.com).



Gambar 11. Kanker kulit ganas (www.kaggle.com).

3.3.2 *Pre-Processing*

Tahap *pre-processing* merupakan tahap persiapan data sebelum pelatihan model dengan melakukan beberapa penyesuaian dan normalisasi terhadap data. Proses data *pre-processing* yang dilakukan pada penelitian ini antara lain normalisasi data dan pembagian data menjadi *data training* dan *data testing*. Normalisasi yang dilakukan pada penelitian ini adalah dengan normalisasi nilai piksel dari rentang 0-255 menjadi rentang 0-1. Pembagian data dilakukan menjadi dua

skema yaitu skema 1 dan skema 2, yang dijelaskan dalam Tabel 3 berikut.

Tabel 3. Skenario pembagian *dataset*

Skema 1		Skema 2	
Ganas	Jinak	Ganas	Jinak
1497	1497	1800 (1497 data asli + 303 data augmentasi)	1800

Pada skema 1 mengikuti jumlah data paling sedikit yaitu data ganas sebanyak 1497 dan mengambil sejumlah data yang sama untuk tahi lalat jinak sehingga total data yang digunakan sebanyak 2994 data. Pada skema 2 data mengikuti jumlah data terbanyak pada data jinak sebanyak 1800 kemudian data kanker ganas diperbanyak dari 1497 menjadi 1800 sehingga total data berjumlah 3600 data. Selanjutnya pembagian *dataset*, pada penelitian ini *dataset* dibagi dengan *k-fold cross-validation*, di mana *k* adalah jumlah *fold* yang akan digunakan. Pada penelitian ini akan menggunakan *k* sebanyak 5 *fold*. Selain itu, untuk *preprocessing data* pada SVM yang harus dilakukan adalah ekstraksi fitur yang relevan dari citra kanker kulit, seperti tekstur, bentuk, atau warna, menggunakan metode seperti *color histogram* atau GLCM (*Gray-Level Co-occurrence Matrix*).

3.3.3 SVM

a. Ekstraksi Fitur

Ekstraksi fitur yang dilakukan pada penelitian ini adalah ekstraksi fitur warna menggunakan, fitur bentuk, serta fitur tekstur. Ekstraksi tekstur dapat menggunakan *Gray-Level Co-occurrence Matrix* (GLCM) dan LBP (*Local Binary Pattern*), ekstraksi warna dapat menggunakan *Color Histogram* dan ekstraksi fitur bentuk berbasis kontur.

b. *Modeling SVM*

Pada SVM proses klasifikasi dilakukan dengan mencari *hyperplane* terbaik untuk memisahkan data antar kelas. Pemodelan untuk klasifikasi SVM dilakukan dengan pemilihan kernel penyesuaian parameter seperti bobot atau koefisien agar yang sesuai dengan model SVM yang akan dibuat. Beberapa jenis *kernel* yang umum digunakan adalah *linear*, RBF (*Radial Basis Function*), dan *polynomial*. Penelitian ini akan menggunakan linear kernel karena cocok dengan data yang kecil dan waktu pelatihan yang singkat.

3.3.4 CNN

a. *Image Augmentation*

Dalam CNN agar menghasilkan akurasi yang tinggi diperlukan jumlah data yang besar untuk proses pelatihan. Hal ini akan menjadi kendala jika ketersediaan *dataset* yang kurang memadai (Sasongko et al., 2023). *Overfitting* merupakan suatu kondisi yang terjadi pada model disebabkan ketersediaan *dataset* yang terbatas sehingga jaringan model *Convolutional Neural Network* tidak dapat mempelajari banyak fitur yang dibutuhkan (Sasongko et al., 2023). Untuk itu proses augmentasi diperlukan untuk mengatasi masalah tersebut. Pendekatan augmentasi data digunakan untuk mengurangi ketergantungan pada data pelatihan dan untuk meningkatkan kinerja model *machine learning* dengan beberapa teknik seperti *flipping*, *rotating*, dan lain sebagainya untuk menambahkan data gambar tanpa mendistorsi data asli (Maharana et al., 2022).

Proses augmentasi data pada CNN yang pertama dengan menggunakan objek `ImageDataGenerator()` yang berisi beberapa parameter seperti *rescale* untuk normalisasi nilai piksel ke rentang 0 dan 1 untuk mempercepat konvergensi selama proses *training*. *Rotation range* dilakukan untuk memberikan variasi gambar dengan

merotasi, *horizontal flip* untuk menciptakan versi vertikal dari gambar, *share range* untuk memberikan pergeseran pada gambar baik vertikal maupun horizontal, dan *fill mode* untuk mengisi piksel yang mungkin kosong akibat proses augmentasi. Selanjutnya menggunakan fungsi objek `flow_from_directory()`, yang memerlukan data untuk dipisahkan ke dalam beberapa direktori untuk keperluan pelatihan dan validasi model.

b. *Modeling*

Pada *Convolutional Neural Network* arsitektur yang digunakan dalam penelitian ini adalah arsitektur *Resnet* ataupun *DenseNet*. Pemilihan kedua arsitektur ini didasarkan pada penelitian sebelumnya dalam (Agarwal & Singh, 2022) dimana dari 4 model yang diujikan model yang menghasilkan tingkat akurasi tertinggi adalah model dengan arsitektur *Resnet50* dan *DenseNet201*. *DenseNet* membantu menyelesaikan masalah *vanishing-gradients* dan memungkinkan penggunaan ulang dan penyebaran fitur yang hilang (Agarwal & Singh, 2022). *ResNet50* memiliki fitur koneksi lompatan, dimana input ke suatu lapisan dapat segera diteruskan ke beberapa lapisan lain, meningkatkan efisiensi jaringan saraf dibandingkan dengan arsitektur lain (Agarwal & Singh, 2022). Selain itu, fitur ini juga membantu mengatasi masalah gradien yang hilang dengan memanfaatkan pemetaan identitas (Agarwal & Singh, 2022).

3.3.5 Evaluasi

Evaluasi ini dilakukan untuk mengukur kinerja dari model yang telah dibangun dalam tugasnya mengklasifikasikan kanker kulit ganas (*malignant*) dan jinak (*benign*). Hasil prediksi harus diuji pada data yang telah dilatih untuk memeriksa secara akurat bagaimana mesin dapat membuat keputusan ini (Nahr et al., 2021). Pada penelitian ini kinerja

model SVM dan CNN dievaluasi menggunakan *confusion matrix*. Parameter yang digunakan di antaranya adalah *accuracy*, *precision*, *recall*, dan *f1-score*. Metrik-metrik ini memberikan gambaran seberapa efektif model SVM dan CNN dapat mengklasifikasikan kanker kulit ganas dan jinak.

Selain itu, pada penelitian ini juga akan membandingkan hasil kinerja dari metode *Support Vector Machine* dan *Convolutional Neural Network* dengan beberapa metode *machine learning* pada penelitian sebelumnya yaitu *K-Nearest Neighbor*, *Naive Bayes*, dan *Random Forest*.

V. PENUTUP

5.1 Simpulan

Dari hasil penelitian yang telah dilakukan, dapat ditarik beberapa kesimpulan utama sebagai berikut.

- a. Model *convolutional neural network* berhasil dibangun dengan arsitektur terbaik yaitu arsitektur DenseNet169 pada data *skema 2* dengan *optimizer SGD*, *learning rate* 0.01 dan *batch size* 32 dan *epoch* 20. Hasil evaluasi menunjukkan bahwa CNN memiliki akurasi yang sangat baik, sebesar 0.8908, menjadikannya sebagai metode yang cukup efektif untuk tugas klasifikasi kanker kulit.
- b. Model SVM juga berhasil dikembangkan dan diuji dalam tugas klasifikasi kanker kulit SVM menunjukkan performa yang baik, dengan menggunakan *linear kernel* pada *skema 2 data*, memperoleh akurasi sebesar 0.8300. Meskipun hasilnya sedikit di bawah performa CNN dalam hal akurasi model SVM tetap memberikan kontribusi signifikan dengan hasil yang cukup konsisten untuk deteksi kanker kulit.
- c. Evaluasi menunjukkan bahwa model CNN memiliki performa yang lebih baik dibandingkan dengan SVM dalam hal *akurasi*, *presisi*, *recall*, dan *f1-score*. Perbandingan dengan penelitian terdahulu juga menunjukkan bahwa CNN merupakan metode yang lebih efektif untuk klasifikasi kanker kulit, mengungguli metode-metode lain yang digunakan sebelumnya.

5.2 Saran

Adapun saran yang diberikan dalam penelitian ini meliputi hal hal sebagai berikut.

- a. Melanjutkan penelitian dengan membandingkan model lain dengan pendekatan *deep learning* yang lebih modern dapat memberikan perspektif tambahan dan mungkin meningkatkan performa dalam klasifikasi kanker kulit.
- b. Pada penelitian selanjutnya dapat memperluas cakupan *dataset* yang digunakan dalam penelitian ini dengan menambahkan lebih banyak variasi gambar, termasuk berbagai jenis kanker kulit dan kondisi pencahayaan.

DAFTAR PUSTAKA

- Almaiah, M. A., Almomani, O., Alsaaidah, A., Al-Otaibi, S., Bani-Hani, N., Hwaitat, A. K. Al, Al-Zahrani, A., Lutfi, A., Awad, A. B., & Aldhyani, T. H. H. (2022). Performance Investigation of Principal Component Analysis for Intrusion Detection System Using Different Support Vector Machine Kernels. *Electronics*, *11*(21), 3571. <https://doi.org/10.3390/electronics11213571>
- Agarwal, K., & Singh, T. (2022). Classification of skin cancer images using convolutional neural networks. *ArXiv Preprint ArXiv:2202.00678*.
- Agustina, R., Magdalena, R., & Pratiwi, N. K. C. (2022). Klasifikasi Kanker Kulit menggunakan Metode Convolutional Neural Network dengan Arsitektur VGG-16. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, *10*(2), 446. <https://doi.org/10.26760/elkomika.v10i2.446>
- Aishwarya, N., Manoj Prabhakaran, K., Debebe, F. T., Reddy, M. S. S. A., & Pranavee, P. (2023). Skin Cancer diagnosis with Yolo Deep Neural Network. *Procedia Computer Science*, *220*, 651–658. <https://doi.org/10.1016/j.procs.2023.03.083>
- Ameri, A. (2020). A Deep Learning Approach to Skin Cancer Detection in Dermoscopy Images. *Journal of Biomedical Physics and Engineering*, *10*(6). <https://doi.org/10.31661/jbpe.v0i0.2004-1107>
- Ayyıldız, N. (2023). *Prediction of Stock Market Index Movements with Machine Learning*. Özgür Yayınları. <https://doi.org/10.58830/ozgur.pub354>
- Caelen, O. (2017). A Bayesian interpretation of the confusion matrix. *Annals of Mathematics and Artificial Intelligence*, *81*(3–4), 429–450. <https://doi.org/10.1007/s10472-017-9564-8>
- Chen, W., Sun, Q., Wang, J., Dong, J.-J., & Xu, C. (2018). A Novel Model Based on AdaBoost and Deep CNN for Vehicle Classification. *IEEE Access*, *6*, 60445–60455. <https://doi.org/10.1109/ACCESS.2018.2875525>
- Chlap, P., Min, H., Vandenberg, N., Dowling, J., Holloway, L., & Haworth, A. (2021). A review of medical image data augmentation techniques for deep

- learning applications. *Journal of Medical Imaging and Radiation Oncology*, 65(5), 545–563. <https://doi.org/10.1111/1754-9485.13261>
- Craythorne, E., & Al-Niami, F. (2017). Skin cancer. *Medicine*, 45(7), 431–434. <https://doi.org/10.1016/j.mpmed.2017.04.003>
- Faruk, M., & Nafi'iyah, N. (2020). Klasifikasi Kanker Kulit Berdasarkan Fitur Tekstur, Fitur Warna Citra Menggunakan SVM dan KNN. *Telematika*, 13(2), 100–109.
- Gomes, R., Kamrowski, C., Langlois, J., Rozario, P., Dircks, I., Grottodden, K., Martinez, M., Tee, W. Z., Sargeant, K., LaFleur, C., & Haley, M. (2022). A Comprehensive Review of Machine Learning Used to Combat COVID-19. *Diagnostics*, 12(8), 1853. <https://doi.org/10.3390/diagnostics12081853>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>
- Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2016). *Densely Connected Convolutional Networks*.
- Javaid, M., Haleem, A., Pratap Singh, R., Suman, R., & Rab, S. (2022). Significance of machine learning in healthcare: Features, pillars and applications. *International Journal of Intelligent Networks*, 3, 58–73. <https://doi.org/10.1016/j.ijin.2022.05.002>
- Khasanah, N., Komarudin, R., Afni, N., Maulana, Y. I., & Salim, A. (2021). Skin Cancer Classification Using Random Forest Algorithm. *SISFOTENIKA*, 11(2), 137. <https://doi.org/10.30700/jst.v11i2.1122>
- Khater, T., Ansari, S., Mahmoud, S., Hussain, A., & Tawfik, H. (2023). Skin cancer classification using explainable artificial intelligence on pre-extracted image features. *Intelligent Systems with Applications*, 20, 200275. <https://doi.org/10.1016/j.iswa.2023.200275>
- Kurniadi, B. W., Prasetyo, H., Ahmad, G. L., & Bagas Aditya, W. (2021). Analisis Perbandingan Algoritma SVM dan CNN untuk Klasifikasi Buah. *Seminar Nasional Mahasiswa Ilmu Komputer Dan Aplikasinya (SENAMIKA)*, 1–11.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lyu, Z., Yu, Y., Samali, B., Rashidi, M., Mohammadi, M., Nguyen, T. N., & Nguyen, A. (2022). Back-Propagation Neural Network Optimized by K-Fold Cross-Validation for Prediction of Torsional Strength of Reinforced Concrete Beam. *Materials*, 15(4), 1477. <https://doi.org/10.3390/ma15041477>
- Mathew, A., Amudha, P., & Sivakumari, S. (2021). *Deep Learning Techniques: An Overview* (pp. 599–608). https://doi.org/10.1007/978-981-15-3383-9_54

- Mishra, R. K., Reddy, G. Y. S., & Pathak, H. (2021). The Understanding of Deep Learning: A Comprehensive Review. *Mathematical Problems in Engineering*, 2021, 1–15. <https://doi.org/10.1155/2021/5548884>
- Mohan, L., Pant, J., Suyal, P., & Kumar, A. (2020). Support Vector Machine Accuracy Improvement with Classification. *2020 12th International Conference on Computational Intelligence and Communication Networks (CICN)*, 477–481. <https://doi.org/10.1109/CICN49253.2020.9242572>
- Monika, M. K., Arun Vignesh, N., Usha Kumari, Ch., Kumar, M. N. V. S. S., & Lydia, E. L. (2020). Skin cancer detection and classification using machine learning. *Materials Today: Proceedings*, 33, 4266–4270. <https://doi.org/10.1016/j.matpr.2020.07.366>
- Murugan, A., Nair, S. A. H., & Kumar, K. P. S. (2019). Detection of Skin Cancer Using SVM, Random Forest and kNN Classifiers. *Journal of Medical Systems*, 43(8), 269. <https://doi.org/10.1007/s10916-019-1400-8>
- Nahr, J. G., Nozari, H., & Sadeghi, M. E. (2021). Artificial intelligence and Machine Learning for Real-world problems (A survey). *International Journal of Innovation in Engineering*, 1(3), 38–47. <https://doi.org/10.59615/ijie.1.3.38>
- Oktora, M. Z., Haiga, Y., & Liana, N. (2023). Peningkatan Pengetahuan Masyarakat terhadap Pengaruh Paparan Sinar Matahari yang dapat Menyebabkan Kanker Kulit. *Jurnal Pengabdian Masyarakat Kesehatan (JURABDIKES)*, 1(2), 28–31. <https://doi.org/10.56260/jurabdikes.v1i2.105>
- Savera, T. R., Suryawan, W. H., & Setiawan, A. W. (2020). Deteksi Dini Kanker Kulit menggunakan K-NN dan Convolutional Neural Network. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 7(2), 373–378. <https://doi.org/10.25126/jtiik.2020702602>
- Sharma, S., Sharma, S., & Athaiya, A. (2020). ACTIVATION FUNCTIONS IN NEURAL NETWORKS. *International Journal of Engineering Applied Sciences and Technology*, 04(12), 310–316. <https://doi.org/10.33564/IJEAST.2020.v04i12.054>
- Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1), 60. <https://doi.org/10.1186/s40537-019-0197-0>
- Sundaram, J., Gowri, K., Devaraju, S., Gokuldev, S., Jayaprakash, S., Anandaram, H., Manivasagan, C., & Thenmozhi, M. (2023). *An Exploration of Python Libraries in Machine Learning Models for Data Science* (pp. 1–31). <https://doi.org/10.4018/978-1-6684-8696-2.ch001>
- Telaumbanua, F. D., Hulu, P., Nadeak, T. Z., Lumbantong, R. R., & Dharma, A. (2019). Penggunaan Machine Learning Di Bidang Kesehatan. *Jurnal Teknologi Dan Ilmu Komputer Prima (JUTIKOMP)*, 2(2), 391–399.

- Van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). scikit-image: image processing in Python. *PeerJ*, 2, e453. <https://doi.org/10.7717/peerj.453>
- Waheed, Z., Waheed, A., Zafar, M., & Riaz, F. (2017). An efficient machine learning approach for the detection of melanoma using dermoscopic images. *2017 International Conference on Communication, Computing and Digital Systems (C-CODE)*, 316–319. <https://doi.org/10.1109/C-CODE.2017.7918949>
- White, J., & Power, S. D. (2023). k-Fold Cross-Validation Can Significantly Over-Estimate True Classification Accuracy in Common EEG-Based Passive BCI Experimental Designs: An Empirical Investigation. *Sensors*, 23(13), 6077. <https://doi.org/10.3390/s23136077>
- Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, 9(4), 611–629. <https://doi.org/10.1007/s13244-018-0639-9>