

Implementasi Jaringan Saraf Konvolusional dengan Inception-V3 untuk Deteksi Katarak Menggunakan Gambar Digital Funduskopi

Muhammad Ahnaf Amrullah, dan Mohammad Isa Irawan
Departemen Matematika, Institut Teknologi Sepuluh Nopember (ITS)
E-Mail: mii@its.ac.id

Abstrak—Katarak merupakan salah satu penyakit mata yang paling serius yang dapat menyebabkan kebutaan. Deteksi dan pengobatan dini dapat mengurangi kebutaan pada pasien katarak. Seiring berkembangnya teknologi pelayanan kesehatan saat ini mengintegrasikan alat kesehatan dan teknologi informasi untuk meningkatkan kualitas dan produktivitas dalam pelayanan kesehatan. Hasil gambar funduskopi atau gambar bagian belakang dan dalam mata (*fundus*) dapat digunakan untuk memprediksi katarak. Dalam Penelitian ini diimplementasikan *Convolutional Neural Network (CNN)* dengan arsitektur *Inception-V3* dalam deteksi katarak berdasarkan gambar digital funduskopi. Terdapat 3 jenis citra fundus yang digunakan yaitu citra fundus normal, citra fundus katarak, dan citra fundus degenerasi makula. Data gambar fundus dipraproses menggunakan *histogram equalization* dan *Contrast Limited Adaptive Histogram Equalization (CLAHE)* terhadap *channel* hijau. Hasil terbaik pada Penelitian ini adalah model dengan praproses CLAHE dengan *Fine Tuning* yang memiliki akurasi sebesar 98,33%.

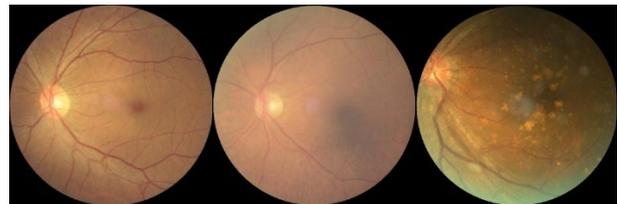
Kata Kunci—*Convolutional Neural Network (CNN)*, *Inception-V3*, Katarak.

I. PENDAHULUAN

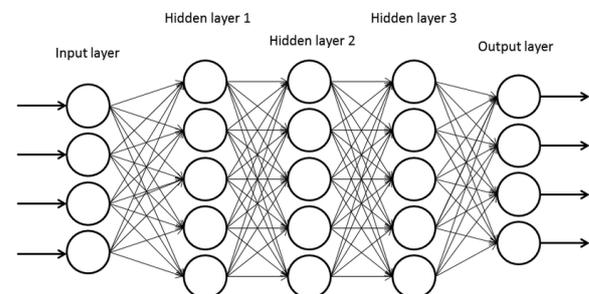
MANUSIA memiliki 5 indera penting dalam tubuhnya, mata yang sering disebut sebagai panca indera. Mata merupakan salah satu dari alat indera yaitu indera penglihatan. Mata merupakan alat indera terpenting bagi manusia karena melalui mata manusia menerima lebih dari 80% informasi visual yang digunakan dalam berbagai aktivitas. Namun tak sedikit orang yang mengalami gangguan pada penglihatan, gangguan penglihatan bervariasi mulai dari yang ringan hingga berat bahkan berpotensi menimbulkan kebutaan.

Katarak merupakan penyebab utama dari kebutaan dan gangguan penglihatan di seluruh dunia. Pada tahun 2020 terdapat sebanyak 15,2 juta kasus katarak yang menyebabkan kebutaan terhadap manusia diatas 50 tahun, menjadi angka tertinggi penyebab kebutaan. Dan sebanyak 78,8 juta kasus katarak menyebabkan gangguan penglihatan terhadap manusia diatas 50 tahun, menjadi angka tertinggi kedua penyebab gangguan penglihatan [1]. Kondisi ini paling umum terjadi pada populasi dengan status sosial ekonomi rendah dan negara berkembang [2].

Pada tahun 2014-2016 dilakukan *Rapid Assessment of Avoidable Blindness (RAAB)* di 15 provinsi Di Indonesia (Nusa Tenggara Barat, Sulawesi Selatan, Jawa Barat, Jawa Timur, Jawa Tengah, Bali, Jakarta, Sumatera Selatan, Kalimantan Selatan, Sulawesi Utara, Sumatera Barat, Nusa Tenggara Timur, Sumatera Utara, Papua Barat, dan Maluku). RAAB adalah metode untuk pengumpulan data kebutaan dan gangguan penglihatan terhadap penduduk berusia diatas 50



Gambar 1. Gambar digital funduskopi normal, katarak, dan AMD.



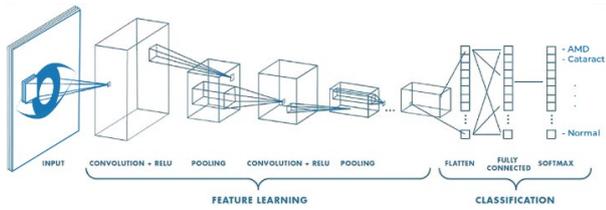
Gambar 2. Deep learning neural network.

tahun yang direkomendasikan oleh WHO. Hasil dari RAAB tersebut menunjukkan angka kebutaan di Indonesia sebesar 929711 dan 720727 kebutaan disebabkan oleh katarak [3].

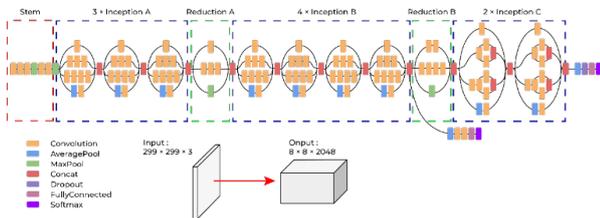
Seiring dengan berkembangnya teknologi, pelayanan kesehatan saat ini mengintegrasikan alat kesehatan dan teknologi informasi untuk meningkatkan kualitas dan produktivitas dalam pelayanan kesehatan. Sistem pelayanan kesehatan berbasis komputer menjadi solusi di daerah berkembang dimana sumber daya medis langka. Hasil gambar funduskopi atau gambar bagian belakang dan dalam mata (*fundus*) dapat disimpan secara digital. Gambar digital funduskopi dapat digunakan untuk deteksi katarak.

Pada bidang teknologi banyak berkembang teori-teori yang berhubungan dengan pemrograman, seperti *Convolutional Neural Network (CNN)* yang merupakan salah satu jenis algoritma dari *deep learning* yang dapat digunakan untuk mengolah data citra. Seiring dengan berjalannya waktu metode CNN juga mengalami perkembangan, dimana saat ini sudah terdapat berbagai macam arsitektur CNN, salah satunya adalah *Inception-V3*. *Inception-V3* merupakan arsitektur CNN yang mendapatkan *1st runner up* dalam *ImageNet Large Scale Visual Recognition Challenge (ILSVRC)* tahun 2015. ILSVRC merupakan kompetisi tahunan yang melombakan klasifikasi terhadap data *ImageNet* yang berisi sebanyak 1,2 juta citra dengan 1000 kelas.

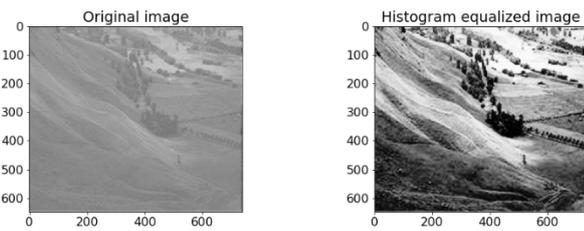
Salah satu penelitian dengan menggunakan model *Inception-V3* terhadap citra medis pernah dilakukan oleh Ahmet Demir dan Feyza Yilmaz terhadap gambar kanker kulit [4]. Untuk data non medis *Inception-V3* pernah



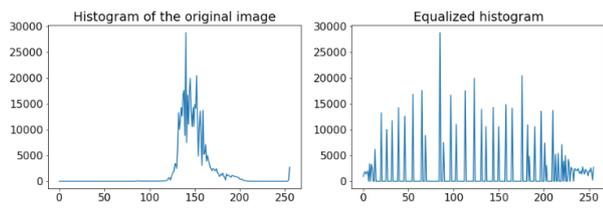
Gambar 3. Contoh arsitektur CNN.



Gambar 4. Arsitektur *inception-V3*.



Gambar 5. Contoh citra sebelum dan setelah histogram *equalization*.



Gambar 6. Histogram citra sebelum dan setelah histogram *equalization*.



Gambar 7. Original histogram dan *clipped* histogram.

digunakan dalam penelitian yang dilakukan oleh Xiaoling Xia et al. Penelitian ini menggunakan model berbasis *Inception-V3* untuk mengklasifikasi 2 dataset bunga. Masing-masing dataset adalah *Oxford-17 flower dataset* dan *Oxford-102 flower dataset* [5].

II. TINJAUAN PUSTAKA

A. Katarak

Katarak merupakan penyebab utama dari gangguan penglihatan di seluruh dunia. Usia lanjut merupakan faktor utama penyebab katarak [6]. Selain usia ada beberapa faktor yang juga mampu menyebabkan katarak yaitu : faktor keturunan, infeksi karena trauma, peradangan, gangguan metabolik seperti diabetes, gangguan nutrisi, merokok, alkohol, obat, dan sinar ultraviolet [7].

Penglihatan penderita katarak akan terasa seperti melihat kabut yang dapat mengganggu aktivitas atau pekerjaan sehari-hari. Lensa mata manusia terdiri atas air dan protein. Katarak terbentuk ketika terjadi penumpukan protein di lensa sehingga dapat mencegah atau mengurangi cahaya yang masuk melewati lensa ke retina. Katarak juga merupakan

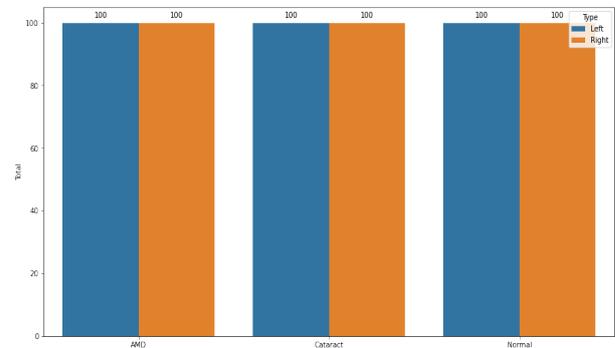
Predicted Class

		Predicted Class	
		0 (Negative)	1 (Positive)
Actual Class	0 (Negative)	True Negative (TN)	False Positive (FP)
	1 (Positive)	False Negative (FN)	True Positive (TP)

Gambar 8. *Confusion matrix* 2 kelas.

		Nilai True Positive			Nilai False Negative			Nilai False Positive		
		Predicted Class			Predicted Class			Predicted Class		
Actual Class	A	a	b	c	a	b	c	a	b	c
	C	d	e	f	d	e	f	d	e	f
	N	g	h	i	g	h	i	g	h	i
		TP(A) = a TP(C) = e	TP(N) = i	FN(A) = b+c FN(C) = d+f	TP(N) = g+h	FP(A) = d+g FP(C) = b+h	TP(N) = c+f			

Gambar 9. Menentukan TP, FN, FP untuk 3 kelas.

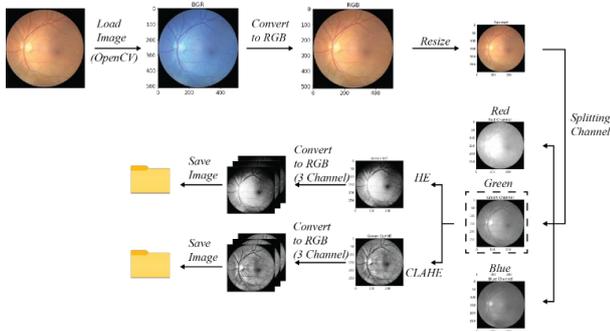


Gambar 10. Jumlah data AMD, katarak, dan normal yang akan digunakan.

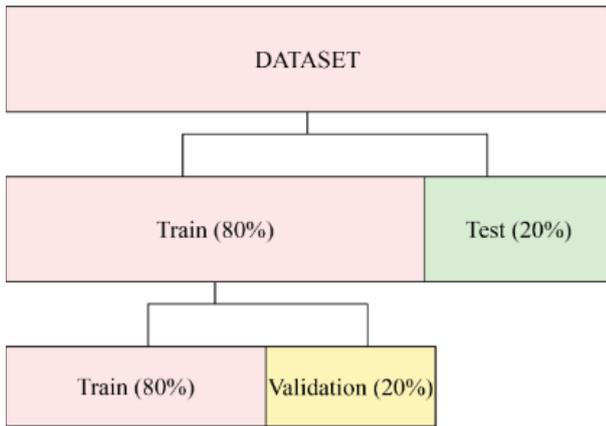
salah satu penyakit mata yang menyebabkan kebutaan.

B. Oftalmoskopi

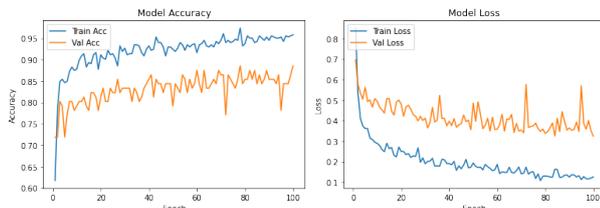
Oftalmoskopi atau funduskopi merupakan bagian dari pemeriksaan mata atau sering disebut sebagai pemeriksaan retina. Oftalmoskopi merupakan tes yang dilakukan untuk pemeriksaan bagian belakang dan dalam mata (*fundus*). Bagian tersebut termasuk retina, cakram optik, dan pembuluh darah. Dalam pemeriksaan oftalmoskopi tentunya diperlukan sebuah alat, alat yang digunakan untuk pemeriksaan oftalmoskopi yaitu oftalmoskop. Oftalmoskop adalah alat seperti senter dengan beberapa lensa kecil yang dapat menunjukkan bagian dalam bola mata. Dengan oftalmoskop, dokter mampu mengidentifikasi masalah mata dan kemungkinan penyakit lainnya. Gambar digital funduskopi dapat dilihat pada Gambar 1.



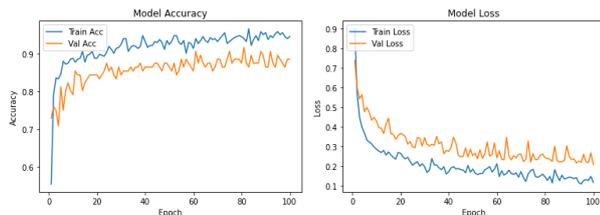
Gambar 11. Alur preprocessing image.



Gambar 12. Pembagian data untuk train set, validation set, dan test set.



Gambar 13. Hasil accuracy, validation accuracy, loss dan validation loss terbaik percobaan 2.



Gambar 14. Hasil accuracy, validation accuracy, loss dan validation loss terbaik percobaan 1.

C. Data Mining

Istilah *Deep Learning* diperkenalkan oleh Geoffrey Hinton pada tahun 2006. Geoffrey Hinton mempublikasikan sebuah paper berjudul “A fast learning algorithm for deep belief nets” yang memperkenalkan salah satu varian dari *Artificial Neural Network* (ANN) yang bernama *deep belief nets* [8]. Pada saat itu juga *Deep Learning* dikenal mampu melakukan training model ANN dengan banyak *layer*.

Deep Learning merupakan bagian dari *Artificial Intelligence* dan *Machine Learning* yang berbasis ANN dengan *hidden layer* yang banyak serta mampu untuk mempelajari fitur data secara otomatis. Gambar 2 merupakan contoh arsitektur dari *Deep Learning*. Meskipun *Deep Learning* menjadi terkenal, ketertarikan terhadap deep

Tabel 1.

Arsitektur inception-V3 yang digunakan

Nama Layer	Input	Output Size
StemBlock	(299x299x3)	(35x35x192)
Modul A-1	(35x35x192)	(35x35x256)
Modul A-2	(35x35x256)	(35x35x288)
Modul A-3	(35x35x288)	(35x35x288)
Reduction-A	(35x35x288)	(17x17x768)
Modul B-1	(17x17x768)	(17x17x768)
Modul B-2	(17x17x768)	(17x17x768)
Modul B-3	(17x17x768)	(17x17x768)
Modul B-4	(17x17x768)	(17x17x768)
Reduction-B	(17x17x768)	(8x8x1280)
Modul C-1	(8x8x1280)	(8x8x2048)
Modul C-2	(8x8x2048)	(8x8x2048)
Global Average Pooling	(8x8x2048)	(1x1x2048)
Dense	(1x1x2048)	(1x1x2048)
Dropout (0.2)	(1x1x2048)	(1x1x2048)
Dense	(1x1x2048)	(1x1x2048)
Softmax	(1x1x2048)	(1x1x3)

Tabel 2.

Hasil accuracy, validation accuracy, loss dan validation loss percobaan 1

Epoch	Batch Size	Train Accuracy	Validation Accuracy	Train Loss	Validation Loss
25	8	0,9115	0,8333	0,2039	0,4372
	16	0,9063	0,8229	0,2507	0,4481
	32	0,8880	0,8229	0,3083	0,4891
	64	0,8698	0,8229	0,3640	0,5412
50	8	0,9531	0,8437	0,1407	0,3944
	16	0,9167	0,8229	0,2018	0,4457
	32	0,9036	0,8020	0,2527	0,4540
	64	0,8828	0,8229	0,2992	0,4879
75	8	0,9635	0,8758	0,1335	0,3604
	16	0,9063	0,7916	0,2193	0,4370
	32	0,9245	0,8333	0,2238	0,4441
	64	0,9010	0,8333	0,2668	0,4686
100	8	0,9661	0,8854	0,1030	0,3249
	16	0,9505	0,8645	0,1480	0,3856
	32	0,9297	0,8229	0,1909	0,3979
	64	0,8984	0,8020	0,2392	0,4699

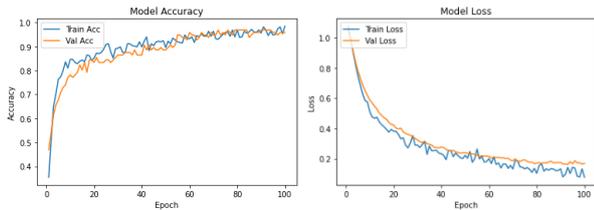
Tabel 3.

Hasil accuracy, validation accuracy, loss dan validation loss percobaan 2

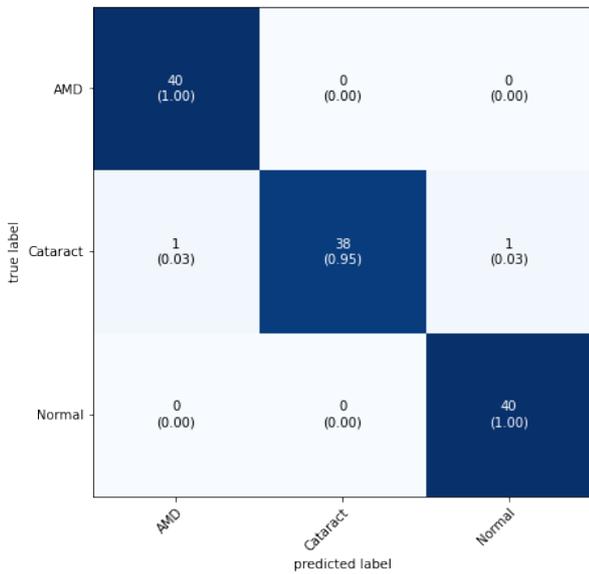
Epoch	Batch Size	Train Accuracy	Validation Accuracy	Train Loss	Validation Loss
25	8	0,9297	0,8750	0,1738	0,2895
	16	0,9010	0,8125	0,2488	0,3923
	32	0,8958	0,8020	0,3084	0,4469
	64	0,8802	0,7500	0,3697	0,5412
50	8	0,9453	0,8854	0,1579	0,2577
	16	0,9271	0,8646	0,1997	0,3061
	32	0,8958	0,8229	0,2635	0,4006
	64	0,8854	0,8125	0,3050	0,4573
75	8	0,9557	0,8750	0,1414	0,2325
	16	0,9349	0,8333	0,1825	0,3209
	32	0,8984	0,8438	0,2307	0,3259
	64	0,9063	0,8333	0,2594	0,4045
100	8	0,9765	0,8854	0,1015	0,2091
	16	0,9401	0,8750	0,1620	0,2574
	32	0,9323	0,8542	0,1933	0,2960
	64	0,9036	0,8438	0,2353	0,3625

learning masih kurang dikarenakan ANN dengan banyak layer memiliki beban komputasi yang besar. Dimana teknologi pada saat itu belum memadai sehingga dibutuhkan komputer dengan spesifikasi yang tinggi.

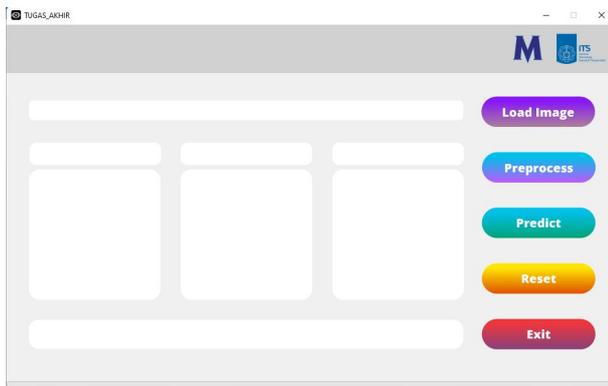
Pada tahun 2009 Andrew ng melalui papernya yang berjudul “Large-scale Deep Unsupervised Learning using Graphics Processors” memperkenalkan penggunaan *Graphics Processing Unit* (GPU) untuk deep learning [9]. Andre ng pada saat itu menggunakan GPU Nvidia GeForce GTX 280. Penggunaan GPU tersebut membuat *Artificial*



Gambar 15. Hasil accuracy, validation accuracy, loss dan validation loss percobaan 3.



Gambar 16. Confusion matrix percobaan 3.



Gambar 17. Tampilan awal GUI.

Neural Network (ANN) berjalan lebih cepat dibandingkan dengan penggunaan CPU. Seiring berkembangnya teknologi yang semakin pesat, penggunaan *Deep Learning* semakin ramai. Dikarenakan *hardware* yang digunakan semakin canggih dan memadai. Banyak hasil dari penggunaan *Deep Learning* yang dapat dinikmati seperti *self-driving car*, pengenalan wajah, pengenalan suara, dan lain-lain.

D. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan hasil dari pengembangan *Multi-Layer Perceptron* (MLP) yang dirancang untuk mengolah data berupa gambar. CNN juga dikenal sebagai *ConvNets*, merupakan salah satu algoritma *Deep Learning*.

CNN pertama kali dikembangkan oleh Kunihiko Fukushima dengan nama *NeoCognitron* pada tahun 1980 [10]. Pada tahun 1989 dikembangkan kembali oleh Yann LeChun dengan nama *LeNet*. Penelitian ini mengenai pengenalan angka kode pos dengan tulisan tangan [11]. Pada

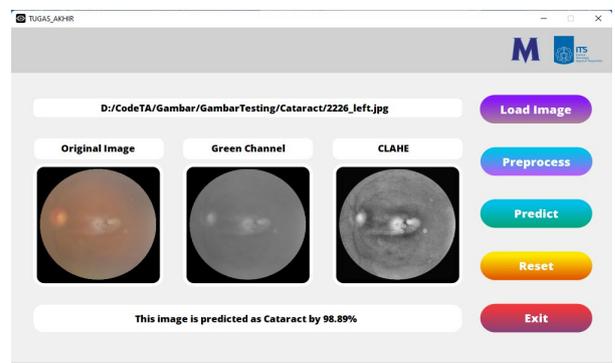
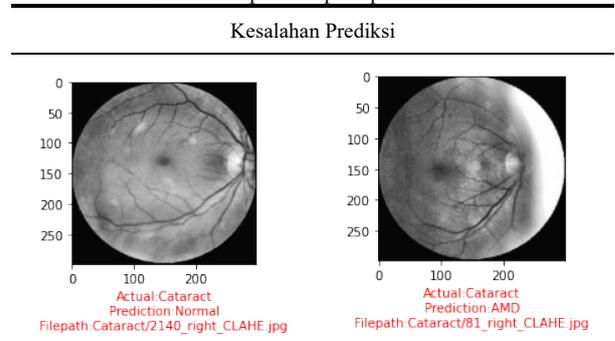
Tabel 4. Hasil accuracy, validation accuracy, loss dan validation loss percobaan 3

Epoch	Train Accuracy	Validation Accuracy	Train Loss	Validation Loss
1	0,3542	0,4688	1,0758	1,0497
2	0,5104	0,5417	0,9888	0,9806
3	0,6510	0,6146	0,9063	0,9125
4	0,7031	0,6562	0,8241	0,8425
5	0,7630	0,7083	0,7526	0,7360
⋮	⋮	⋮	⋮	⋮
96	0,9505	0,9583	0,1374	0,1659
97	0,9740	0,9479	0,0834	0,1868
98	0,9818	0,9583	0,0784	0,1732
99	0,9541	0,9583	0,1329	0,1638
100	0,9844	0,9583	0,0767	0,1682

Tabel 5. Hasil precision, recall, dan f1-score percobaan 3

	Precision	Recall	F1-Score
AMD	0.9756	1	0.9877
Cataract	1	0.95	0.9743
Normal	0.9756	1	0.9877

Tabel 6. Kesalahan prediksi pada percobaan 3



Gambar 18. Tampilan prediksi GUI.

tahun 2012, Alex Krizhevsky dengan CNN miliknya yang bernama *AlexNet* berkompetisi di *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) pada 30 September 2012 dan berhasil menjadi juara dalam kompetisi tersebut [12]. *AlexNet* mendapatkan *top-5 error rate* sebesar 15,3%, presentase ini 10,9% lebih baik dari juara 2 yaitu sebesar 26,2%. Sejak saat itu, pemanfaatan CNN meroket sekaligus pembuktian bahwa metode *Deep Learning*, khususnya CNN mampu mengungguli metode *Machine Learning* terhadap kasus klasifikasi objek pada citra. Gambar 3 merupakan contoh arsitektur dari CNN.

E. Inception-V3

GoogLeNet merupakan CNN yang diusulkan oleh Google pada tahun 2014. *GoogLeNet* mengadopsi struktur jaringan *Inception*, yang tidak hanya mengurangi jumlah parameter

jaringan, tetapi juga meningkatkan kedalaman jaringan. Oleh karena itu, *GoogLeNet* banyak digunakan dalam tugas klasifikasi gambar. Karena inti dari jaringan *GoogLeNet* adalah struktur jaringan *Inception*, jaringan *GoogLeNet* juga disebut jaringan *Inception*.

Inception-V3 merupakan pengembangan dari jaringan *Inception-V1* atau *GoogLeNet* [13]. *Inception-V3* merupakan struktur jaringan yang dikembangkan oleh *Keras*, yang telah dilatih sebelumnya di *ImageNet*. Ukuran input gambar default adalah 299×299 dengan tiga *channel*. Arsitektur jaringan *Inception-V3* ditunjukkan pada Gambar 4.

F. Histogram Equalization

Histogram Equalization (HE) merupakan salah satu teknik yang digunakan untuk meningkatkan gambar atau image enhancement. Histogram Equalization adalah teknik pemrosesan gambar yang menyesuaikan kontras gambar dengan menggunakan histogramnya. Konsep dari Histogram Equalization meratakan persebaran nilai intensitas piksel suatu citra. Untuk meningkatkan kontras gambar, Histogram Equalization menyebarkan nilai intensitas piksel yang paling sering atau merentangkan rentang intensitas gambar [14]. Dengan melakukan ini, pemerataan histogram memungkinkan area gambar dengan kontras yang lebih rendah untuk mendapatkan kontras yang lebih tinggi. Hasil dari Histogram Equalization dapat dilihat pada Gambar 5 dan Gambar 6.

G. Contrast Limited Adaptive Histogram Equalization

Contrast Limited Adaptive Histogram Equalization (CLAHE) merupakan sebuah varian dari *Adaptive Histogram Equalization* (AHE). Cara kerja CLAHE berbeda dengan *Histogram Equalization* yang bekerja secara global. Dalam penggunaan CLAHE citra dibagi menjadi beberapa bagian atau blok. Setiap bagian tersebut diterapkan *Histogram Equalization* dan digabungkan kembali. Dalam CLAHE jika histogram diatas atau melebihi *clip limit* yang telah ditentukan, piksel tersebut akan didistribusikan seragam dibawah area *clip limit* seperti pada Gambar 7 [15].

H. Evaluasi Kinerja

Evaluasi kinerja merupakan hal yang dilakukan untuk mengukur keberhasilan suatu model. Pada umumnya dalam mengevaluasi hasil klasifikasi yaitu menggunakan *accuracy*, *precision*, *recall*, dan *F1-score*. Dalam menentukan nilai *accuracy*, *precision*, *recall*, dan *F1-score* dapat dilihat dengan menggunakan *confusion matrix*. *Confusion matrix* memberikan informasi perbandingan antara kelas yang sebenarnya dengan kelas hasil klasifikasi yang telah dilakukan suatu model. *Confusion matrix* dengan 2 kelas yaitu positif dan negatif diperlihatkan pada Gambar 8.

Dalam *confusion matrix* terdapat 4 istilah yaitu: *True Negative* (TN), *True Positive* (TP), *False Negative* (FN), dan *False Positive* (FP). Misalkan terdapat 2 kelas yaitu mata normal dan mata katarak. Lalu mata katarak dilabeli dengan kelas positif dan mata normal dengan kelas negatif. Sehingga *True Negative* (TN) adalah jumlah mata normal yang diprediksi benar sebagai mata normal. *True Positive* (TP) adalah jumlah mata katarak yang diprediksi benar sebagai mata katarak. *False Negative* (FN) adalah jumlah mata katarak yang diprediksi salah sebagai mata normal. *False Positive* (FP) adalah jumlah mata normal yang diprediksi

salah sebagai mata katarak. Dalam menentukan nilai *accuracy*, *precision*, *recall*, dan *f1-score* dapat dilihat dengan menggunakan *confusion matrix* dengan perhitungan yang didefinisikan pada persamaan (1) - (4) [16]:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Dari persamaan di atas *Accuracy* adalah rasio data diprediksi Benar (positif dan negatif) dengan semua data yang diprediksi. *Precision* adalah rasio data positif diprediksi benar dengan semua data yang diprediksi positif. *Recall* adalah rasio data positif diprediksi benar dengan semua data yang positif. Sedangkan *F1 Score* merupakan *harmonic mean* dari *precision* dan *recall*.

Pada penelitian ini digunakan 3 kelas. Kelas yang akan digunakan yaitu mata normal, katarak, dan degenerasi makula atau *Age-related Macular Degeneration* (AMD). Tidak seperti pada klasifikasi biner, tidak ada kelas positif maupun negatif. Sehingga untuk penyelesaian dari 3 kelas harus mencari nilai TP, FP dan FN pada masing-masing kelas. maka untuk *confusion matrix* 3 kelas seperti pada Gambar 9

Nilai *accuracy* pada klasifikasi 3 kelas didapatkan dengan membagi semua data yang terprediksi benar dengan semua data. Dalam menentukan nilai *accuracy* dapat dilihat dengan menggunakan *confusion matrix* pada Gambar 9 dengan perhitungan yang didefinisikan pada persamaan berikut:

$$Accuracy = \frac{TP(A)+TP(B)+TP(C)}{a+b+c+d+e+f+g+h+i} \quad (5)$$

Nilai *precision*, *recall*, dan *f1-score* untuk klasifikasi 3 kelas dihitung tiap kelas n dengan $n = \{A, C, N\}$. Dalam menentukan nilai *precision*, *recall*, dan *f1-score* untuk klasifikasi 3 kelas dapat dilihat dengan menggunakan *confusion matrix* pada Gambar 9 dengan perhitungan yang didefinisikan pada persamaan berikut:

$$Precision(n) = \frac{TP(n)}{TP(n)+FP(n)} \quad (6)$$

$$Recall(n) = \frac{TP(n)}{TP(n)+FN(n)} \quad (7)$$

$$F1\ Score(n) = 2 \times \frac{Precision(n) \times Recall(n)}{Precision(n) + Recall(n)} \quad (8)$$

III. METODE PENELITIAN

A. Pengambilan Data

Pada tahap ini dilakukan pengambilan data yang berupa dataset gambar digital funduskopi. Dataset tersebut berisi gambar digital funduskopi yang telah memiliki label normal dan label memiliki penyakit seperti katarak, diabetes, degenerasi makula, hipertensi, miopia patologis, dan penyakit/kelainan lainnya. Dataset tersebut merupakan gambar digital funduskopi yang telah dikumpulkan oleh Shangong Medical Technology Co., Ltd. dari berbagai

rumah sakit atau pusat kesehatan di China yang disimpan di situs kaggle.

B. Pre-processing Data

Pada tahap ini dilakukan pemilahan data yang berupa dataset gambar digital funduskopi dimana data yang digunakan adalah gambar digital funduskopi dengan label normal, katarak dan degenerasi makula. Data yang telah diseleksi tersebut merupakan data yang akan digunakan dalam penelitian ini diambil 200 data untuk masing-masing kelas sehingga total data yang digunakan sebanyak 600 [17].

Data citra yang telah diseleksi dilakukan *resize* menjadi ukuran 299x299 dimana ukuran tersebut adalah ukuran *default input* dari *Inception-V3*. Citra yang telah dilakukan *resize Histogram Equalization* terhadap channel hijau untuk meningkatkan kontras dari gambar [18]. Selain *Histogram Equalization* akan dilakukan juga percobaan lainnya dengan menerapkan *contrast limited adaptive histogram equalization* (CLAHE) terhadap channel hijau [19]. Data yang telah dilakukan praproses akan dibagi menjadi data latih, dan data uji dengan presentase sebesar 80% dan 20%. Dari data latih yang telah dibagi diambil 20% data untuk dijadikan data validasi.

C. Evaluasi Hasil

Pada tahap ini dilakukan evaluasi dari hasil klasifikasi yang dihasilkan oleh model yang telah dibuat. Tahap evaluasi ini menggunakan *confusion matrix*. Dari *confusion matrix* dapat dihitung *akurasi*, *recall*, *precision*, dan *f1-score*.

D. Pembuatan User Interface (UI)

Pada tahap ini dilakukan pembuatan *Graphical User Interface* (GUI) yang lebih mudah dioperasikan sehingga hasil dari penelitian dapat diterima oleh pembaca. Model *Inception-V3* yang digunakan dalam GUI merupakan model terbaik yang didapat dari tahap sebelumnya.

IV. HASIL DAN PEMBAHASAN

A. Informasi Dataset

Penelitian ini menggunakan dataset yang di peroleh pada situs kaggle. Dataset bernama *Ocular Disease Intelligent Recognition* (ODIR). ODIR merupakan database yang terdiri dari 6.392 gambar digital funduskopi berwarna. Kumpulan gambar *fundus* tersebut dikumpulkan oleh Shangong Medical Technology Co., Ltd. dari berbagai rumah sakit atau pusat medis di Tiongkok. Gambar fundus tersebut diambil menggunakan berbagai kamera seperti Canon, Zeiss dan Kowa.

B. Pre-processing Data

Pada langkah ini dilakukan pemilahan data dimana pada penelitian Penelitian ini akan dilakukan klasifikasi terhadap citra yang berjenis mata normal, katarak, dan degenerasi makula atau *Age-Related Macular Degeneration* (AMD). Sehingga data yang digunakan hanya data dengan label mata AMD, katarak, dan normal. Jumlah data yang digunakan berjumlah 600 dapat dilihat pada Gambar 10.

C. Pre-processing Image

Pada tahap ini dilakukan *preprocessing image* terhadap

dataset. *Preprocessing image* dilakukan agar pelatihan dapat berjalan dengan baik sehingga data yang akan diklasifikasi juga akan terklasifikasi dengan baik juga. Channel hijau gambar dapat meningkatkan kontras antara bagian depan dan pembuluh darah sambil menjaga sebagian besar detail dari gambar asli. Oleh karena itu diterapkan *histogram equalization* terhadap channel hijau dari gambar untuk meningkatkan kontras global [18]. Tahap *preprocessing image* pertama akan dilakukan *splitting channel* RGB. Diterapkan *Histogram Equalization* dan CLAHE terhadap channel hijau dan dicari yang terbaik antara kedua filter tersebut.

Sebelum dilakukan proses *Histogram equalization* maupun CLAHE dilakukan *resize* terhadap data untuk mengurangi beban komputasi. *Resize* dilakukan dengan mengubah ukuran citra yang awalnya berukuran 512 x 512 menjadi 299x299 yaitu ukuran *input default* dari model *Inception-V3*. Setelah dilakukan HE dan CLAHE gambar dikonversi menjadi 3 channel agar bisa digunakan dalam model *Inception-V3* yang memiliki input channel berjumlah 3. Hasil gambar yang telah dilakukan *preprocessing* akan disimpan dalam folder yang akan digunakan untuk tahap selanjutnya. Alur yang terdapat dalam *preprocessing image* dapat dilihat pada Gambar 11.

D. Splitting Data

Gambar yang telah dilakukan *preprocessing* dibagi sebagai data *train* dan *testing* dengan persentase masing-masing sebesar 80% dan 20% dipilih secara acak. Dari data *train* akan dibagi sebagai data *train* dan *validation* dengan persentase masing-masing sebesar 80% dan 20% secara acak. Pembagian dataset dapat dilihat pada Gambar 12.

E. Perancangan Model Inception-V3

Pada tahap ini dilakukan perancangan model dengan arsitektur *Inception-V3*. Dalam penelitian ini dilakukan modifikasi pada bagian *classification*. Karena jumlah kelas dalam penelitian ini berjumlah 3 kelas maka *layer* terakhir model akan diganti dengan 3 kelas. Pada *classification layer* yang baru ditambahkan *Global Average Pooling*, dan *Dropout* sebesar 0.2. Arsitektur utuh model *Inception-V3* dalam penelitian ini dapat dilihat pada Tabel 1.

F. Training Model

1) Percobaan 1

Percobaan pertama dilakukan terhadap data *Histogram Equalization* dengan *learning rate* 0.001, nilai epsilon dengan nilai 0.1, *batch size* sebesar 8, 16, 32, 64, dengan 25, 50, 75 dan 100 *epoch*. Hasil *accuracy*, *validation accuracy*, *loss* dan *validation loss* percobaan 1 dapat dilihat pada Tabel 2. Untuk grafik Hasil *accuracy*, *validation accuracy*, *loss* dan *validation loss* terbaik dalam percobaan 1 dapat dilihat pada Gambar 13.

2) Percobaan 2

Percobaan kedua dilakukan terhadap data CLAHE dengan parameter yang sama dengan percobaan 2. Hasil *accuracy*, *validation accuracy*, *loss* dan *validation loss* percobaan 2 dapat dilihat pada Tabel 3. Untuk grafik Hasil *accuracy*, *validation accuracy*, *loss* dan *validation loss* terbaik dalam percobaan 2 dapat dilihat pada Gambar 14.

3) Percobaan 3

Dari percobaan 1 dan 2 dapat dilihat bahwa pelatihan dataset HE maupun CLAHE dengan parameter *Epoch* 100 dan *batch size* 8 memiliki *validation loss* paling rendah tetapi keduanya masih mengalami *overfitting*. Percobaan 2 atau percobaan dengan menggunakan data CLAHE memiliki *validation loss* lebih rendah sehingga dilakukan percobaan ketiga yaitu percobaan dengan melakukan *Fine Tuning* terhadap dataset CLAHE. Hasil *accuracy*, *validation accuracy*, *loss* dan *validation loss* percobaan 3 dapat dilihat pada Tabel 4. Untuk grafik Hasil *accuracy*, *validation accuracy*, *loss* dan *validation loss* terbaik dalam percobaan 3 dapat dilihat pada Gambar 15.

G. Evaluasi Hasil

Model yang telah dilatih akan dilakukan pengujian terhadap data *test*. Hasil prediksi dari model ditampilkan dalam *confusion matrix*. *Confusion matrix* pada percobaan 3 dapat dilihat pada Gambar 16. Untuk nilai *precision*, *recall*, dan *f1-score* untuk masing-masing kelas pada percobaan 2 dapat dilihat pada Tabel 5. Dari *confusion matrix* pada Gambar 16 terdapat 4 kesalahan prediksi. Kesalahan prediksi tersebut 2 citra katarak yang salah diprediksi menjadi normal dan AMD. Hasil kesalahan prediksi akan dimuat pada Tabel 6.

H. GUI

Setelah mendapatkan model yang terbaik, dibentuk GUI. GUI tersebut akan memiliki fungsi untuk memuat citra, *preprocess* citra hingga untuk memprediksi citra *fundus*. Hasil GUI yang telah dibuat dapat dilihat pada Gambar 17 dan Gambar 18.

V. KESIMPULAN

Beberapa kesimpulan yang dapat diambil dalam penelitian ini adalah (1) Convolutional Neural Network (CNN) dengan arsitektur Inception-V3 dapat digunakan untuk prediksi katarak berdasarkan citra digital funduskopi dengan sangat baik. (2) Dalam Penelitian ini performansi hasil terbaik implementasi model CNN dengan arsitektur Inception-V3 menggunakan preprocessing Contrast Limited Adaptive Histogram Equalization (CLAHE) dengan Fine Tuning. Performansi terbaik ditunjukkan dengan nilai akurasi sebesar 98,33%.

Beberapa hal yang masih dapat ditingkatkan dari penelitian ini adalah sebagai berikut: (1) Diperlukan data yang jauh lebih banyak untuk menghasilkan model yang lebih cerdas. Dalam penelitian ini data untuk penyakit katarak dan AMD terbatas. (2) Perlu dilakukan penelitian lebih lanjut mengenai efek dari resolusi citra terhadap akurasi. Dalam penelitian ini citra yang awalnya memiliki resolusi 512x512 dilakukan resize menjadi 299x299. (3) Dapat dilakukan percobaan lebih lanjut dengan arsitektur Convolutional Neural Network (CNN) yang lebih baru.

DAFTAR PUSTAKA

- [1] G. B. D. Collaborators and V. Study, "Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the right to sight: an analysis for the global burden of disease study," *Lancet Glob Health*, vol. 9, no. 2, pp. 144–160, Feb. 2021, doi: [https://doi.org/10.1016/S2214-109X\(20\)30489-7](https://doi.org/10.1016/S2214-109X(20)30489-7).
- [2] M. Khairallah *et al.*, "Number of people blind or visually impaired by cataract worldwide and in world regions, 1990 to 2010," *Invest Ophthalmol Vis Sci*, vol. 56, no. 11, pp. 6762–6769, Oct. 2015.
- [3] L. Sulistyowati, *Peta Jalan Penanggulangan Gangguan Penglihatan di Indonesia tahun 2017-2030*. Jakarta: Kementerian Kesehatan RI, 2017.
- [4] A. Demir and F. Yilmaz, "Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3," *Medical Technologies Congress (TIPTEKNO)*, pp. 1–4, 2019, doi: [10.1109/iptekno47231.2019.8972045](https://doi.org/10.1109/iptekno47231.2019.8972045).
- [5] X. Xia, C. Xu, and B. Nan, "Inception-V3 for Flower Classification," in *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, 2017, pp. 783–787. doi: [10.1109/ICIVC.2017.7984661](https://doi.org/10.1109/ICIVC.2017.7984661).
- [6] A. M. Cynthia, *Optometric Clinical Practice Guideline Care of the Adult Patient with Optometric Paractical Clinic Guideline*. America: American Optometric Association, 2004.
- [7] American Optometric Association, "Evidence-based clinical practice guideline: comprehensive pediatric eye and vision examination," *Optometric Clinical Practice*, vol. 2, no. 2, pp. 2–67, 2020, doi: [10.37685/uwlibraries.2575-7717.2.2.1007](https://doi.org/10.37685/uwlibraries.2575-7717.2.2.1007).
- [8] G. E. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput*, vol. 18, no. 7, pp. 1527–1554, 2006, doi: <https://doi.org/10.1162/neco.2006.18.7.1527>.
- [9] R. Raina, A. Madhavan, and A. Ng, "Large-Scale Deep Unsupervised Learning Using Graphics Processors," in *International Conference on Machine Learning*, 2009.
- [10] K. Fukushima and S. Miyake, "Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Visual Pattern Recognition," in *Competition and Cooperation in Neural Nets*, 1982, pp. 267–285.
- [11] Y. LeCun *et al.*, "Handwritten Digit Recognition with a Back-Propagation Network," in *Neural Information Processing Systems (NIPS)*, 1989, pp. 396–404.
- [12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun ACM*, vol. 60, pp. 84–90, 2012.
- [13] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception architecture for computer vision," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, 2015, doi: [10.1109/CVPR.2016.308](https://doi.org/10.1109/CVPR.2016.308).
- [14] C. Rafael, *Digital Image Processing*, 4th ed. New York: Pearson, 2018.
- [15] S. M. Pizer, R. E. Johnston, J. P. Ericksen, B. C. Yankaskas, and K. E. Muller, "Contrast-Limited Adaptive Histogram Equalization: Speed and Effectiveness," in *Proceedings of the First Conference on Visualization in Biomedical Computing*, 1990, pp. 337–345. doi: [10.1109/VBC.1990.109340](https://doi.org/10.1109/VBC.1990.109340).
- [16] M. Vakili, M. K. Ghamsari, and M. Rezaei, "Performance analysis and comparison of machine and deep learning algorithms for IoT data classification," *ArXiv:2001.09636*, Jan. 2020, doi: [10.48550/ARXIV.2001.09636](https://doi.org/10.48550/ARXIV.2001.09636).
- [17] T. Pratap and P. Kokil, "Computer-aided diagnosis of cataract using deep transfer learning," *Biomed Signal Process Control*, vol. 53, p. 101533, 2019, doi: <https://doi.org/10.1016/j.bspc.2019.04.010>.
- [18] J.-J. Yang *et al.*, "Exploiting ensemble learning for automatic cataract detection and grading," *Comput Methods Programs Biomed*, vol. 124, pp. 45–57, 2016, doi: <https://doi.org/10.1016/j.cmpb.2015.10.007>.
- [19] S. B. Júnior and D. Welfer, "Automatic detection of microaneurysms and hemorrhages in color eye fundus images," *International Journal of Computer Science and Information Technology*, vol. 5, pp. 21–37, 2013.