



Penerapan Metode CNN (Convolutional Neural Network) untuk Mengklasifikasikan Jenis Cacat pada Kulit Hewan

Eka Legya Frannita^{1,*}, Alifia Revan Prananda²

¹ Program Studi Teknologi Pengolahan Produk Kulit, Politeknik ATK Yogyakarta, Yogyakarta, Indonesia

² Program Studi Teknologi Informasi, Fakultas Teknik, Universitas Tidar, Magelang, Indonesia

Email: ^{1,*}eka.legya@atk.ac.id, ²revan@untidar.ac.id

Email Penulis Korespondensi: eka.legya@atk.ac.id

Abstrak—Industri pembuatan produk fashion dari kulit hewan merupakan salah satu bidang industri yang sedang berkembang. Di Indonesia, industri produk fashion dari kulit hewan merupakan salah satu industri prioritas pemerintah. Hal ini dikarenakan terdapat cukup banyak industri produk fashion dari kulit hewan yang berkembang di berbagai daerah di Indonesia. Pesatnya perkembangan industri ini juga diiringi dengan banyaknya jumlah permintaan konsumen terhadap produk fashion dari kulit hewan. Melihat fakta tersebut, maka penjaga kualitas kulit yang dihasilkan adalah hal yang sangat penting agar kulit yang diproduksi dapat memenuhi kebutuhan konsumen. Salah satu langkah untuk menjaga kualitas kulit adalah dengan melakukan proses inspeksi kualitas kulit. Namun, saat ini proses inspeksi kulit masih dilakukan secara manual dengan mengidentifikasi jenis-jenis cacat yang terdapat pada permukaan kulit secara langsung. Proses pemeriksaan yang manual ini tentunya memiliki kelemahan yang dapat merugikan seperti membutuhkan waktu yang relatif lama, butuh ketelitian yang tinggi, butuh operator yang berpengalaman. Adapun tujuan dari penelitian ini adalah mengembangkan arsitektur convolutional neural network yang dapat mengklasifikasikan jenis-jenis cacat pada kulit guna menyelesaikan permasalahan di atas. Kontribusi dari penelitian ini adalah arsitektur yang dikembangkan bertujuan untuk mengolah data dengan jumlah kelas dan jumlah dataset yang cukup banyak dan belum banyak digunakan oleh penelitian sebelumnya. Penelitian dimulai dengan proses studi literatur dan pengumpulan data. Data yang digunakan dalam penelitian ini merupakan dataset publik yang terdiri dari 3600 citra kulit digital yang terdistribusi ke dalam enam kelas (folding mask, grain off, growth marks, loose grains, pinhole, non-defective). Setelah data terkumpul selanjutnya dilakukan pelatihan dan pengujian. Proses pelatihan dilakukan pada beberapa epoch untuk mencari model dengan performa terbaik. Berdasarkan proses pelatihan dan pengujian model diperoleh akurasi pelatihan sebesar 90,43% dan akurasi pengujian sebesar 88,47%.

Kata Kunci: Convolutional Neural Network; Cacat Kulit; Inspeksi Kulit; Klasifikasi; Kualitas Kulit

Abstract—Recently, leather industry was rapidly growth in several countries. In Indonesia, leather industry became one of the government's priority industries since there were quite a lot of leather industries developing in various regions in Indonesia. On the other hand, there were large number of consumer demand for leather products. Regarding to this fact, maintaining the quality of leather was strongly important. An alternative solution for maintaining leather quality is to conduct leather quality inspection process. However, currently the leather inspection process was still carried out manually by identifying directly the types of defects found on the surface of the leather. This manual inspection process certainly has several hurdles such as time consuming, requiring high accuracy, and requiring experienced operators. This research aimed to develop convolutional neural network architecture that can classify types of leather defects. This research was done by conducting four main processes which were literature study and data collection processes, develop CNN architecture, training process, and testing process. This research work used public dataset consisting of 3600 digital leather images distributed into six classes (folding mask, grain off, growth marks, loose grains, pinhole, non-defective). Based on the training and testing process, the model obtained training accuracy of 90.43% and testing accuracy of 88.47%.

Keywords: Convolutional Neural Network; Classification; Leather Defect; Leather Inspection; Leather Quality

1. PENDAHULUAN

Industri pengolahan kulit saat ini merupakan industri yang cukup besar dan sedang berkembang pesat. Hal ini beriringan dengan kebutuhan akan konsumsi daging hewan. Agar kulit hewan tersebut tidak menjadi limbah, maka banyak industri yang memulai untuk mengolah kulit hewan menjadi komoditas atau produk kulit yang memiliki nilai jual yang tinggi. Pada tahun 2003, industri kulit global telah berhasil memproduksi sekitar 18 miliar feet persegi, dengan perkiraan nilai sebesar 40 miliar US dolar. Negara-negara berkembang kini memproduksi lebih dari 60% kebutuhan kulit di seluruh dunia. Hal tersebut menunjukkan bahwa industri kulit merupakan industri yang cukup esensial dan memiliki peranan penting di berbagai negara (Aslam et al., 2019) (Moganam & Sathia Seelan, 2022)(Wang & Kyong, 2023).

Sebelum kulit dapat diproses sebagai produk manufaktur, kulit harus melewati serangkaian tahap pengolahan. Proses pengolahan kulit umumnya terdiri atas persiapan penyamakan, penyamakan, dan finishing. Setiap tahap pemrosesan ini melibatkan berbagai prosedur tergantung pada jenis bahan baku dan spesifikasi kulit yang diinginkan untuk produk akhirnya. Setelah melewati proses tersebut, selanjutnya dilakukan proses inspeksi yang bertujuan untuk mengidentifikasi kualitas kulit. Proses inspeksi dilakukan dengan mengidentifikasi jenis-jenis cacat yang muncul pada permukaan kulit (Gan et al., 2022). Saat ini dalam industri pengolahan kulit, proses inspeksi dan evaluasi cacat permukaan kulit masih dilakukan secara manual oleh tenaga kerja manusia. Proses ini melibatkan pemeriksaan secara visual untuk mengidentifikasi dan mengkategorikan cacat-cacat seperti goresan, bintik-bintik, atau ketidaksempurnaan lainnya yang mungkin terdapat pada permukaan kulit. Namun, dengan meningkatnya permintaan akan produksi kulit dalam skala besar, terdapat tantangan signifikan terkait efisiensi dan kecepatan dalam proses ini. Proses inspeksi manual cenderung memakan waktu dan membutuhkan banyak tenaga kerja manusia, yang tidak hanya meningkatkan biaya

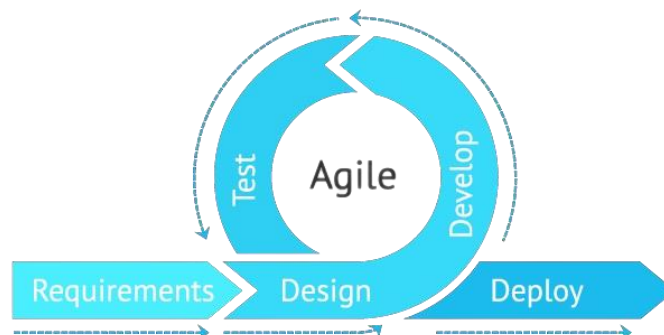
produksi tetapi juga memperlambat laju produksi keseluruhan. Selain itu, keberadaan inspektur manusia dapat menyebabkan variasi subjektif dalam penilaian cacat, yang mengarah pada penilaian yang tidak konsisten dari waktu ke waktu (Moganam & Sathia Seelan, 2022)(Liong, Gan, Huang, Yuan, et al., 2019)(Vasagam & Sornam, 2022)(Aslam et al., 2020)(Deng et al., 2020).

Melihat kondisi proses inspeksi kulit yang masih manual, maka diperlukan suatu inovasi yang dapat diterapkan untuk proses inspeksi agar dapat berjalan lebih efektif dan efisien. Saat ini, pemanfaatan artificial intelligence telah banyak dikembangkan dan dimanfaatkan untuk berbagai bidang. Sebagai contoh, saat ini mulai banyak dikembangkan sistem cerdas pada bidang kedokteran, telekomunikasi, keuangan, diagnosis medis, transportasi, pengambilan informasi, energi, dan sebagainya. Dalam bidang industri kulit, pemanfaatan artificial intelligence telah dilakukan oleh beberapa peneliti terdahulu (Vasagam & Sornam, 2022)(Abdullah et al., 2024)(Liong, Gan, Huang, Liu, et al., 2019). Sebagai contoh adalah penelitian yang dilakukan oleh Shahzaib Iqbal dkk tentang pengembangan metode MLR-Net yang merupakan metode multi-layer residual convolutional network untuk melakukan segmentasi pada objek cacat kulit. Penelitian ini menggunakan dataset citra kulit digital. Penelitian ini berhasil memperoleh sensitivitas sebesar 87.50%, spesifisitas sebesar 99.78%, akurasi sebesar 95.98%, F1-score sebesar 85.90%, dan Intersection over Union (IoU) sebesar 78.98% (Iqbal et al., 2023). Penelitian yang diusulkan oleh Salik dkk tentang pengembangan metode semantic segmentation untuk mendeteksi cacat pada permukaan kulit. Penelitian ini menggunakan dataset MVTEC pada proses eksperimen metode. Hasil dari penelitian ini adalah penelitian ini berhasil mendapatkan Intersection over Union (IoU) sebesar 78.98% sebesar 94% (Khanal et al., 2022). Penelitian yang dilakukan oleh Siyamalan dkk tentang pengembangan metode kolaborasi deep semi-supervised learning dengan knowledge distillation untuk mengklasifikasikan cacat pada permukaan kulit. Penelitian ini berhasil mendapatkan akurasi klasifikasi sebesar 97% dengan menggunakan dataset SLS (Manivannan, 2023). Penelitian yang dilakukan oleh Azween dkk tentang pengembangan metode klasifikasi cacat kulit menggunakan Mask Region-based Convolution Neural Network Model. Penelitian ini berhasil memperoleh rata-rata akurasi sebesar lebih dari 90% (Abdullah et al., 2024).

Penelitian-penelitian di atas pada dasarnya telah menghasilkan akurasi yang cukup tinggi. Namun, beberapa penelitian di atas hanya fokus untuk mengklasifikasi beberapa jenis cacat kulit saja. Penelitian ini bertujuan untuk mengklasifikasikan jenis-jenis cacat kulit dengan variasi yang lebih banyak. Adapun jenis cacat dalam penelitian ini adalah folding mask, grain off, growth marks, loose grains, pinhole, non-defective yang mana merupakan cacat kulit yang sering muncul pada saat pengolahan kulit di industri. Dalam penelitian ini dilakukan eksperimen dengan beberapa model setting pada arsitektur CNN untuk mendapatkan model dengan performa terbaik.

2. METODOLOGI PENELITIAN

Penelitian ini menerapkan metode penelitian agile yang terdiri dari beberapa tahapan utama yakni tahap identifikasi masalah dan identifikasi kebutuhan, tahap perancangan, tahap pengembangan, tahap pengujian dan terakhir tahap deployment seperti yang digambarkan pada Gambar 1. Pada penelitian ini, tahap identifikasi masalah dan identifikasi kebutuhan bertujuan untuk mencari literatur arsitektur deep learning yang sesuai dengan kasus. Tahap perancangan merupakan tahap perancangan arsitektur deep learning. Tahap pengembangan atau develop merupakan tahap yang bertujuan untuk mengembangkan arsitektur deep learning. Tahap pengujian atau tahap test merupakan tahap yang bertujuan untuk menguji hasil model deep learning yang sudah dikembangkan. Tahap deployment adalah tahap launching hasil running arsitektur deep learning. Dalam penelitian ini, tahap deployment merupakan tahap yang digunakan untuk mengeksport model hasil training dan testing yang dilakukan (Molla et al., 2023).



Gambar 1. Agile method (Molla et al., 2023)

2.1 Kerangka Dasar Penelitian

Dalam penelitian ini dilakukan pengembangan model CNN yang memiliki kemampuan untuk mengklasifikasikan jenis-jenis cacat yang terdapat pada permukaan kulit. Adapun deskripsi kerangka dasar penelitian pada penelitian ini adalah sebagai berikut:

- a. Tujuan penelitian

Penelitian ini memiliki tujuan untuk mengembangkan model deep learning dengan menggunakan arsitektur convolutional neural network yang dapat membedakan karakteristik dari masing-masing jenis cacat yang terdapat pada kulit.

b. Scope atau batasan penelitian

Batasan model convolutional neural network yang dikembangkan adalah model convolutional neural network yang memiliki kemampuan untuk mengklasifikasikan objek dalam bentuk citra digital. Adapun batasan atau scope data yang digunakan adalah penelitian ini hanya menggunakan data citra kulit tersamak.

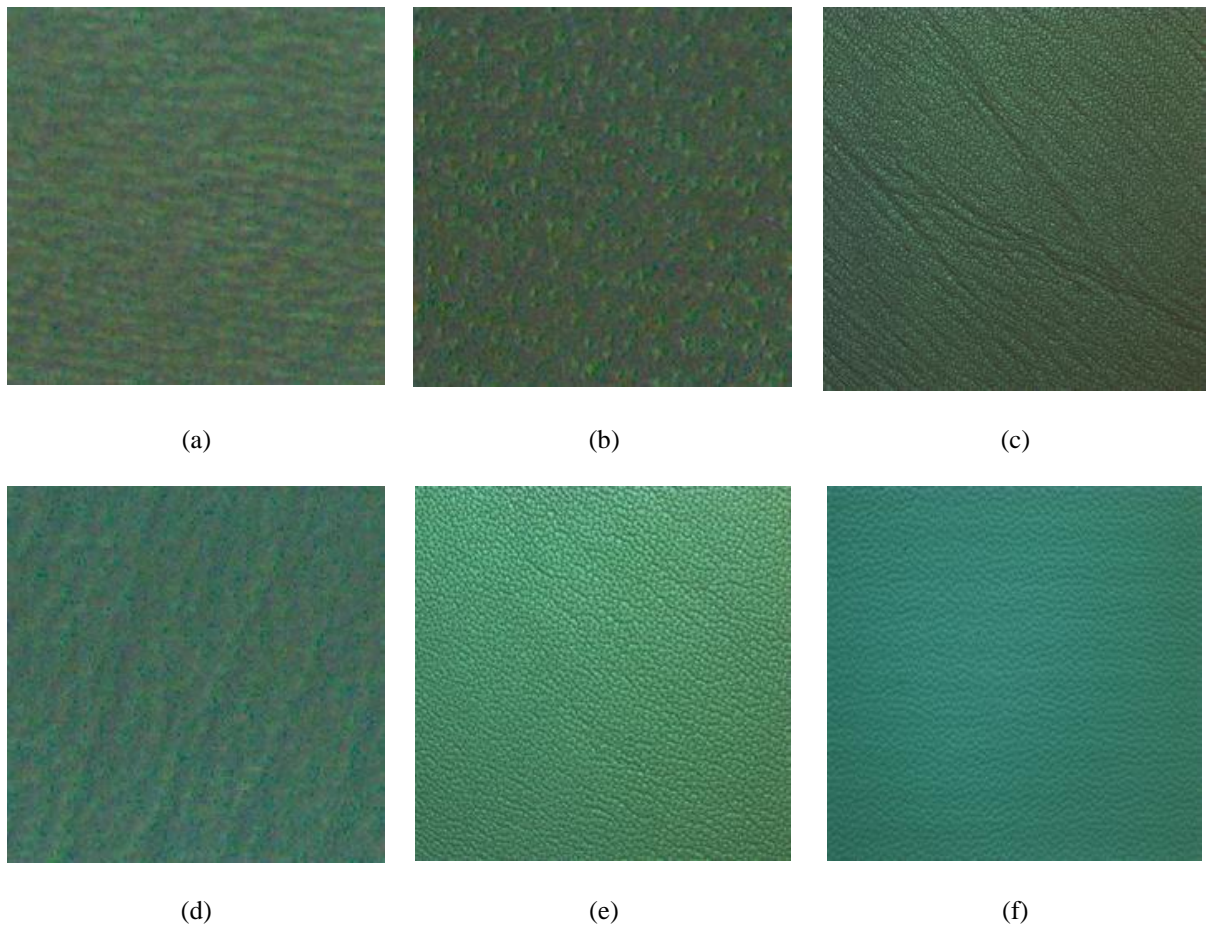
c. Data yang digunakan

Penelitian ini menggunakan dataset publik atau dataset yang terdapat pada repository umum. Tabel 1 menunjukkan deskripsi dataset yang digunakan.

Tabel 1. Deskripsi data yang digunakan

Item	Penjelasan
Tipe data	Leather defect images
Ukuran data	227 x 227 piksel
Ground truth	Klasifikasi yang terdiri dari enam kelas yakni folding mask, grain off, growth marks, loose grains, pinhole, non-defective.
Proporsi data	Dataset terdiri atas 3600 citra dengan proporsi masing-masing kelas terdiri atas 600 citra.
Referensi data	Dataset publik yang dikembangkan oleh Moganam dkk (Moganam & Sathia Seelan, 2022)(Praveen Kumar & Denis Ashok, 2020)(Moganam & Seelan, 2020)

Adapun contoh dataset yang digunakan dalam penelitian ini diilustrasikan dalam Gambar 2. Dalam ilustrasi gambar tersebut terdapat enam gambar yang mengilustrasikan masing-masing kelas yakni kelas folding mask, grain off, growth marks, loose grains, pinhole, non-defective.



Gambar 2. Contoh dataset yang digunakan dalam penelitian dengan rincian sebagai berikut: (a) folding mask, (b) grain off, (c) growth marks, (d) loose grains, (e) pinhole, dan (f) non-defective

d. Hipotesis



TIN: Terapan Informatika Nusantara

Vol 5, No 2, July 2024, page 125-134

ISSN 2722-7987 (Media Online)

Website <https://ejournal.seminar-id.com/index.php/tin>

DOI 10.47065/tin.v5i2.5390

Hipotesis dalam penelitian ini adalah model convolutional neural network yang diusulkan dapat mengklasifikasikan jenis-jenis cacat pada kulit dengan baik.

2.2 Tahapan Penelitian

Tahapan dalam penelitian ini terdiri dari tiga tahapan utama, yakni tahap pengembangan arsitektur CNN, tahap pelatihan model dan tahap pengujian model. Adapun gambaran tahapan penelitian dapat dilihat pada Gambar 3.



Gambar 3. Tahapan penelitian

- Tahap studi literatur dan pengumpulan data
Pada tahap ini dilakukan studi literatur untuk menganalisis bagaimana tren penggunaan arsitektur deep learning pada kasus klasifikasi cacat kulit atau pada kasus serupa. Pada tahap ini juga dilakukan proses pengumpulan data yang sesuai dengan kasus.
- Tahap perancangan arsitektur CNN
Arsitektur CNN yang dikembangkan dalam penelitian ini tentunya adalah arsitektur yang sesuai untuk kasus multi-classification dengan jumlah data yang banyak. Berikut adalah arsitektur yang digunakan dalam penelitian ini:

Tabel 2. Arsitektur deep learning yang digunakan

Layer Name	Layer Type	nIn, nOut	Total Parameters	Shape	Layer Input
in	Input layer	-, -	-	-	-
0	Convolution Layer	1, 20	520	W:{20,1,5,5}, b:{1,20}	[in]
1	Activation Layer	-, -	0	-	[0]
2	Subsampling Layer	-, -	0	-	[1]
3	Convolution Layer	2, 50	25, 050	W:{50,20,5,5}, b:{1,50}	[2]
4	Activation Layer	-, -	0	-	[3]
5	Subsampling Layer	-, -	0	-	[4]
6	Dense Layer	2450, 500	1, 225, 500	W:{2450, 500}, b:{1,500}	[5]
7	Activation Layer	-, -	0	-	[6]
	Output Layer	500, 6	3, 006	W:{500, 6}, b:{1, 6}	[7]

- Tahap pelatihan model
Proses pelatihan diawali dengan proses pembagian data menjadi 3 bagian dengan proporsi 80:10:10. Proporsi tersebut dapat diartikan bahwa 80% dari total data digunakan sebagai data pelatihan, 10% data digunakan sebagai data validasi dan 10% data digunakan untuk proses pengujian. Data pelatihan dan data validasi digunakan pada saat proses pelatihan data. Pada proses tersebut, data validasi berfungsi sebagai data uji untuk melihat performa pada saat proses pelatihan. Dalam proses pelatihan ini dilakukan beberapa setting yakni proses pelatihan dilakukan pada beberapa epoch meliputi 25 epoch, 50 epoch, 75 epoch, 100 epoch, 125 epoch, 150 epoch, dan 175 epoch. Hal ini dimaksudkan sebagai bentuk eksperimen untuk mencari model dengan performa terbaik.
- Tahap pengujian model
Tahap ini merupakan tahapan yang berfungsi untuk menguji reliabilitas performa model deep learning yang dikembangkan. Dalam proses ini terdapat beberapa evaluation matrix yang digunakan meliputi akurasi, recall, precision, ROC area dan F1-score.

3. HASIL DAN PEMBAHASAN

Penelitian ini fokus pada proses pengembangan arsitektur CNN dengan luaran akhir model deep learning untuk kasus klasifikasi enam kelas. Adapun bentuk pemanfaatan akhir dari model deep learning ini adalah dapat digunakan sebagai bagian dari proses inspeksi kulit yang otomatis. Pengembangan model deep learning ini sangat penting dalam industri kulit karena dapat membantu proses inspeksi kulit menjadi lebih cepat, lebih akurat, dan lebih efektif. Kebutuhan akan tenaga cross-checker pun dapat dibantu dengan adanya model dari hasil penelitian ini. Berikut adalah pemaparan hasil penelitian serta pembahasan dari hasil yang telah diperoleh.



3.1 Hasil proses pelatihan dan pengujian

Proses pelatihan diawali dengan proses pembagian data menjadi tiga proporsi yakni 80% untuk data training atau data pelatihan, 10% untuk data validation dan 10% untuk data testing atau data pengujian. Tabel 3 mendeskripsikan jumlah data pada masing-masing data.

Tabel 3. Pembagian proporsi data

Kelas data	Jumlah data training	Jumlah data validation	Jumlah data testing
folding mask	480	60	60
grain off	480	60	60
growth marks	480	60	60
loose grains	480	60	60
pinhole	480	60	60
non-defective	480	60	60
Total data	2880	360	360

Tabel 4 hasil mendeskripsikan hasil pelatihan atau training dengan 25 epoch, 50 epoch, 75 epoch, 100 epoch, 125 epoch, 150 epoch, dan 175 epoch. Dalam Tabel 4 terdapat beberapa indikator pelatihan yang meliputi true positive rate, false positive rate, precision, recall, F1-score dan ROC area. Semua komponen tersebut digunakan untuk mengevaluasi kinerja deep learning.

Tabel 4. Rekap hasil pelatihan pada masing-masing epoch

Komponen evaluasi	Kelas Data					
	Folding mask	Grain off	Growth marks	Loose grains	Non-defective	Pinhole
25 epoch						
TP rate	0.931	0.529	0.981	0.848	0.711	0.638
FP rate	0.135	0.036	0.056	0.022	0.015	0.011
Precision	0.579	0.751	0.776	0.887	0.908	0.922
Recall	0.931	0.529	0.981	0.848	0.711	0.638
F1-Score	0.714	0.620	0.867	0.867	0.798	0.754
ROC area						
50 epoch						
TP rate	0,867	0,771	0,985	0,933	0,911	0,565
FP rate	0,091	0,062	0,014	0,014	0,010	0,002
Precision	0,655	0,716	0,935	0,929	0,946	0,978
Recall	0,867	0,771	0,985	0,933	0,911	0,565
F1-Score	0,746	0,743	0,959	0,931	0,929	0,716
ROC area	0,963	0,969	0,999	0,996	0,996	0,974
75 epoch						
TP rate	0.919	0.682	0.975	0.950	0.961	0.648
FP rate	0.105	0.034	0.002	0.021	0.007	0.005
Precision	0.635	0.805	0.992	0.899	0.965	0.963
Recall	0.919	0.682	0.975	0.950	0.961	0.648
F1-Score	0.751	0.738	0.983	0.924	0.963	0.775
ROC area	0.971	0.968	1.000	0.996	0.999	0.971
100 epoch						
TP rate	0.854	0.841	0.994	0.971	0.969	0.694
FP rate	0.063	0.059	0.002	0.003	0.003	0.005
Precision	0.728	0.744	0.992	0.983	0.985	0.962
Recall	0.854	0.841	0.994	0.971	0.969	0.694
F1-Score	0.786	0.789	0.993	0.977	0.977	0.806
ROC area	0.972	0.972	1.000	0.999	1.000	0.977
125 epoch						
TP rate	0.906	0.833	0.998	1.000	0.988	0.702
FP rate	0.070	0.039	0.000	0.002	0.000	0.004
Precision	0.721	0.814	0.998	0.990	1.000	0.971
Recall	0.906	0.833	0.998	1.000	0.988	0.702
F1-Score	0.803	0.823	0.998	0.995	0.994	0.815
ROC area	0.984	0.985	1.000	1.000	1.000	0.984
150 epoch						
TP rate	0.844	0.822	0.973	0.858	0.819	0.656
FP rate	0.038	0.099	0.040	0.010	0.009	0.010

Komponen evaluasi	Kelas Data					
	Folding mask	Grain off	Growth marks	Loose grains	Non-defective	Pinhole
Precision	0.815	0.629	0.829	0.943	0.950	0.929
Recall	0.844	0.822	0.973	0.858	0.819	0.656
F1-Score	0.829	0.713	0.895	0.899	0.879	0.769
ROC area	0.980	0.950	0.994	0.986	0.994	0.974
175 epoch						
TP rate	0.900	0.790	0.992	0.927	0.753	0.660
FP rate	0.079	0.083	0.022	0.003	0.001	0.007
Precision	0.693	0.660	0.898	0.982	0.992	0.946
Recall	0.900	0.790	0.992	0.927	0.753	0.660
F1-Score	0.783	0.719	0.943	0.954	0.856	0.778
ROC area	0.977	0.961	0.998	0.998	0.988	0.977

Tabel 5 mendeskripsikan hasil evaluasi berdasarkan waktu training dan validation serta jumlah parameter dan akurasi dari masing-masing epoch. Dalam Tabel 5, dapat diketahui bahwa semakin banyak jumlah epoch tentu akan meningkatkan durasi training. Namun jumlah epoch yang tinggi tidak selalu dapat menghasilkan akurasi yang tinggi.

Tabel 5. Rekap hasil proses training

Jumlah epoch	Durasi training (detik)	Durasi validation (detik)	Total Parameters	Akurasi (%)
25 epoch	198,5	7,39	1.254.076	77,20
50 epoch	391,6	7,31	1.254.076	83,87
75 epoch	586,2	7,28	1.254.076	85,53
100 epoch	776,4	7,30	1.254.076	88,71
125 epoch	972,2	6,86	1.254.076	90,43
150 epoch	1167,6	6,87	1.254.076	82,87
175 epoch	1344,7	6,92	1.254.076	83,66

Gambar 4 menunjukkan hasil confusion matrix dari masing-masing epoch. Dalam hasil tersebut terlihat perbedaan yang cukup signifikan dari hasil masing-masing epoch. Gambar 4(c) merupakan hasil dengan menggunakan 125 epoch. Pada gambar tersebut terlihat bahwa jumlah true positif atau jumlah yang terdeteksi benar sesuai dengan groundh sangat banyak dibandingkan dengan hasil pada epoch yang lain. Jumlah data yang salah klasifikasi juga relative lebih sedikit dibandingkan dengan hasil pada epoch yang lain.

```

=== Confusion Matrix ===
  a  b  c  d  e  f  <-- classified as
447  1  6  9  1 16 | a = Folding_marks
165 259 12 33 16  5 | b = Grain_off
  0  3 471  1  3  2 | c = Growth_marks
  1 41 13 407 15  3 | d = loose_grains
  0 38 94  8 345  0 | e = non_defective
159  3 11  1  0 306 | f = pinhole
    
```

(a)

```

=== Confusion Matrix ===
  a  b  c  d  e  f  <-- classified as
416 53  0  4  1  6 | a = Folding_marks
 76 378  5 24  7  0 | b = Grain_off
  0  0 473  0  7  0 | c = Growth_marks
  0 19  5 448  8  0 | d = loose_grains
  0 20 17  6 442  0 | e = non_defective
143 58  6  0  2 271 | f = pinhole
    
```

(b)

```

=== Confusion Matrix ===
  a  b  c  d  e  f  <-- classified as
441 26  0  2  1 10 | a = Folding_marks
111 334  2 38  5  0 | b = Grain_off
  0  0 468  2  8  2 | c = Growth_marks
  1 19  1 456  3  0 | d = loose_grains
  0  9  1  9 466  0 | e = non_defective
142 27  0  0  0 311 | f = pinhole
    
```

(c)

```

=== Confusion Matrix ===
  a  b  c  d  e  f  <-- classified as
410 56  0  0  1 13 | a = Folding_marks
 70 412  0  4  4  0 | b = Grain_off
  0  1 477  2  0  0 | c = Growth_marks
  0 12  0 466  2  0 | d = loose_grains
  0 10  3  2 470  0 | e = non_defective
 83 63  1  0  0 333 | f = pinhole
    
```

(d)

```

=== Confusion Matrix ===
  a  b  c  d  e  f  <-- classified as
435 33  0  2  0 10 | a = Folding_marks
 81 408  0  1  0  0 | b = Grain_off
  0  1 479  0  0  0 | c = Growth_marks
  0  0  0 480  0  0 | d = loose_grains
  0  3  1  2 479  0 | e = non_defective
 87 56  0  0  0 337 | f = pinhole
  
```

(e)

```

=== Confusion Matrix ===
  a  b  c  d  e  f  <-- classified as
405 60  0  4  1 10 | a = Folding_marks
 40 403 14 15 10  8 | b = Grain_off
  0  6 467  1  4  2 | c = Growth_marks
  1 23 35 412  6  3 | d = loose_grains
  0 46 36  5 397  1 | e = non_defective
 51 103 11  0  0 315 | f = pinhole
  
```

(f)

```

=== Confusion Matrix ===
  a  b  c  d  e  f  <-- classified as
432 37  1  2  0  8 | a = Folding_marks
 88 387  9  5  0  1 | b = Grain_off
  0  0 476  1  1  2 | c = Growth_marks
  0 12 19 445  1  3 | d = loose_grains
  0 93 23  0 365  4 | e = non_defective
103 57  2  0  1 317 | f = pinhole
  
```

(g)

Gambar 4. Confusion matrix pada masing-masing epoch dengan rincian : (a) 25 epoch, (b) 50 epoch, (c) 75 epoch, (d) 100 epoch, (e) 125 epoch, (f) 150 epoch, dan (g) 175 epoch

Berdasarkan hasil pada Tabel 4 dan Tabel 5 selanjutnya dipilih model dengan performa terbaik. Dari kedua tabel tersebut dapat disimpulkan bahwa model dengan 125 epoch memiliki performa yang paling baik, sehingga model tersebut digunakan pada proses pengujian atau testing. Gambar 5 menunjukkan hasil dari pengujian.

```

=== Confusion Matrix ===
  a  b  c  d  e  f  <-- classified as
 61 59  0  0  0  0 | a = Folding_marks
  0 120  0  0  0  0 | b = Grain_off
  0  0 115  3  0  2 | c = Growth_marks
  0  0  0 115  5  0 | d = loose_grains
  0  0  0  1 119  0 | e = non_defective
  0 13  0  0  0 107 | f = pinhole
  
```

Gambar 5. Confusion matrix pada hasil testing

Tabel 6. Rekap hasil pelatihan pada masing-masing epoch

Komponen evaluasi	Kelas Data					
	Folding mask	Grain off	Growth marks	Loose grains	Non-defective	Pinhole
TP rate	0.508	1.000	0.958	0.958	0.992	0.892
FP rate	0.000	0.120	0.000	0.007	0.008	0.003
Precision	1.000	0.625	1.000	0.966	0.960	0.982
Recall	0.508	1.000	0.958	0.958	0.992	0.892
F1-Score	0.674	0.769	0.979	0.962	0.975	0.934
ROC area	0.983	0.979	1.000	0.999	1.000	0.985

Hasil akurasi yang diperoleh pada proses testing adalah sebesar 88,47%. Akurasi tersebut menunjukkan bahwa performa model deep learning yang dikembangkan sangat baik karena perbedaan akurasi training dan testing tidak terlalu jauh.

3.2 Pembahasan

Berdasarkan hasil training atau pelatihan serta testing atau pengujian dapat disimpulkan bahwa model pada 125 epoch memiliki performa yang lebih baik dari model yang lain. Pada dasarnya dataset yang digunakan merupakan dataset



publik yang sudah digunakan oleh beberapa peneliti terdahulu. Tabel 7 menunjukkan hasil perbandingan akurasi dari metode yang diusulkan dengan penelitian terdahulu.

Tabel 7. Tabel perbandingan dengan penelitian terdahulu

Peneliti	Metode/arsitektur deep learning	Akurasi testing
(Aslam et al., 2019)	AlexNet	87,60%
	VGG 16	85,80%
	ResNet 50	83,60%
	Inception V3	80,00%
Proposed method	CNN	88,47%

Pada dasarnya, untuk hasil dari penelitian terdahulu cukup baik. Dalam proses training, masing-masing metode yakni AlexNet, VGG 16, ResNet 50 dan Inception V3 mendapatkan akurasi training lebih dari 90%. Namun, pada saat testing terjadi penurunan akurasi yang cukup signifikan hampir lebih dari 10%. Hal ini dapat dikarenakan beberapa hal seperti overfitting. Sedangkan untuk arsitektur yang diusulkan tidak mengalami penurunan akurasi yang drastis. Oleh karena itu dapat disimpulkan bahwa arsitektur yang dikembangkan sangat baik untuk mengklasifikasikan jenis-jenis cacat kulit.

4. KESIMPULAN

Penelitian ini bertujuan untuk mengembangkan arsitektur convolutional neural network untuk mengklasifikasikan jenis-jenis cacat kulit yang proses identifikasi jenis-jenis cacat kulit merupakan tahap penting dalam proses inspeksi kulit sebelum kulit diolah menjadi produk. Data yang digunakan dalam penelitian ini adalah citra digital kulit hewan yang terdiri dari 3600 citra. Data yang digunakan terdistribusi dalam enam kelas yakni folding mask, grain off, growth marks, loose grains, pinhole, dan non-defective yang merupakan jenis-jenis cacat dalam kulit. Masing-masing data terdiri dari 600 citra. Adapun tahapan penelitian yang dilakukan terdiri atas empat tahap utama. Keempat tahap tersebut adalah tahap studi literatur dan pengumpulan data, tahap perancangan arsitektur deep learning, tahap pelatihan model, dan tahap pengujian model. Pada tahap pelatihan model data dibagi menjadi tiga proporsi yakni data training, validation dan testing dengan proporsi 80:10:10. Proses pelatihan dilakukan pada 25 epoch, 50 epoch, 75 epoch, 100 epoch, 125 epoch, 150 epoch, dan 175 epoch. Hasil terbaik selanjutnya dilanjutkan pada proses testing. Berdasarkan proses pelatihan dan pengujian model diperoleh akurasi pelatihan sebesar 90,43% dan akurasi pengujian sebesar 88,47%. Hasil tersebut kemudian dibandingkan dengan penelitian terdahulu. Berdasarkan hasil perbandingan diperoleh informasi bahwa akurasi pengujian yang dihasilkan oleh metode yang diusulkan lebih tinggi daripada penelitian terdahulu, sehingga dapat disimpulkan bahwa metode yang diusulkan memiliki performa yang sangat baik dalam mengklasifikasikan citra digital kulit.

REFERENCES

- Abdullah, A. Bin, Jawahar, M., Manogaran, N., Subbiah, G., Seeranagan, K., Balusamy, B., & Saravanan, A. C. (2024). Leather Image Quality Classification and Defect Detection System using Mask Region-based Convolution Neural Network Model. *International Journal of Advanced Computer Science and Applications*, 15(4), 526–536. <https://doi.org/10.14569/IJACSA.2024.0150455>
- Aslam, M., Khan, T. M., Naqvi, S. S., Holmes, G., & Naffa, R. (2020). Ensemble Convolutional Neural Networks With Knowledge Transfer for Leather Defect Classification in Industrial Settings. *IEEE Access*, 8, 198600–198614. <https://doi.org/10.1109/ACCESS.2020.3034731>
- Aslam, M., Khan, T., Naqvi, S., Holmes, G., & Naffa, R. (2019). On the Application of Automated Machine Vision for Leather Defect Inspection and Grading: A Survey. *IEEE Access*, PP. <https://doi.org/10.1109/ACCESS.2019.2957427>
- Deng, J., Liu, J., Wu, C., Zhong, T., Gu, G., & Ling, B. W.-K. (2020). A Novel Framework for Classifying Leather Surface Defects Based on a Parameter Optimized Residual Network. *IEEE Access*, 8, 192109–192118. <https://doi.org/10.1109/ACCESS.2020.3032164>
- Gan, Y. S., Yau, W. C., Liong, S. T., & Chen, C. C. (2022). Automated Classification System for Tick-Bite Defect on Leather. *Mathematical Problems in Engineering*, 2022. <https://doi.org/10.1155/2022/5549879>
- Iqbal, S., Khan, T. M., Naqvi, S. S., & Holmes, G. (2023). MLR-Net: A multi-layer residual convolutional neural network for leather defect segmentation. *Engineering Applications of Artificial Intelligence*, 126, 107007. <https://doi.org/https://doi.org/10.1016/j.engappai.2023.107007>
- Khanal, S. R., Silva, J., Magalhães, L., Soares, J., Gonzalez, D. G., Castilla, Y. C., & Ferreira, M. J. (2022). Leather Defect Detection Using Semantic Segmentation: A Hardware platform and software prototype. *Procedia Computer Science*, 204, 573–580. <https://doi.org/https://doi.org/10.1016/j.procs.2022.08.070>
- Liong, S.-T., Gan, Y. S., Huang, Y.-C., Liu, K.-H., & Yau, W.-C. (2019). *Integrated Neural Network and Machine Vision Approach For Leather Defect Classification*.



- Liong, S.-T., Gan, Y. S., Huang, Y.-C., Yuan, C.-A., & Chang, H.-C. (2019). *Automatic Defect Segmentation on Leather with Deep Learning*.
- Manivannan, S. (2023). Collaborative deep semi-supervised learning with knowledge distillation for surface defect classification. *Computers & Industrial Engineering*, 186, 109766. <https://doi.org/https://doi.org/10.1016/j.cie.2023.109766>
- Moganam, P. K., & Sathia Seelan, D. A. (2022). Deep learning and machine learning neural network approaches for multi class leather texture defect classification and segmentation. *Journal of Leather Science and Engineering*, 4(1), 7. <https://doi.org/10.1186/s42825-022-00080-9>
- Moganam, P. K., & Seelan, D. A. S. (2020). Perceptron neural network based machine learning approaches for leather defect detection and classification. *Instrumentation Measure Metrologie*, 19(6), 421–429. <https://doi.org/10.18280/I2M.190603>
- Molla, Y. S., Yimer, S. T., & Alemneh, E. (2023). COSMIC-Functional Size Classification of Agile Software Development: Deep Learning Approach. *2023 International Conference on Information and Communication Technology for Development for Africa (ICT4DA)*, 155–159. <https://doi.org/10.1109/ICT4DA59526.2023.10302232>
- Praveen Kumar, M., & Denis Ashok, S. (2020). A multi-level colour thresholding based segmentation approach for improved identification of the defective region in leather surfaces. *Engineering Journal*, 24(2), 101–108. <https://doi.org/10.4186/ej.2020.24.2.101>
- Vasagam, S. N., & Sornam, M. (2022). *Intermittent Leather Defect Detection Based on Ensemble Algorithms Derived from Black Hat Transformation and Hough Transformation BT - ICT Analysis and Applications* (S. Fong, N. Dey, & A. Joshi, Eds.; pp. 35–45). Springer Nature Singapore.
- Wang, F., & Kyoung, K. J. (2023). Leather Defect Detection Method in Clothing Design Based on TDENet. *IEEE Access*, 11, 104890–104904. <https://doi.org/10.1109/ACCESS.2023.3308493>