

Pap Smear Image Classification for Cervical Cancer Prediction with Transfer Learning on ResNet101 Architecture

Sila Cahya Dewi^{1*}, Rumini^{2*}

* Informatika, Universitas AMIKOM Yogyakarta
silacahyadewi@students.amikom.ac.id¹, rumini@amikom.ac.id²

Article Info

Article history:

Received 2025-07-21

Revised 2025-08-21

Accepted 2025-09-03

Keyword:

Deep Learning,
ResNet101,
Transfer Learning,
Feature Extraction,
Cervical Cancer.

ABSTRACT

Early detection of cervical cancer remains a pivotal strategy to improve clinical outcomes and mitigate mortality associated with this disease. This study introduces a robust deep learning framework employing the ResNet101 architecture to facilitate the automated classification of cervical cell images derived from the Pap smear examinations. By leveraging transfer learning, the pre-trained ResNet101 model was fine-tuned to extract salient morphological features critical for distinguishing among diverse cervical cell categories. A comprehensive dataset of labeled Pap smear images, systematically expanded through augmentation techniques, was utilized to enhance model generalizability. The proposed approach achieved a remarkable classification accuracy of 99.7%, highlighting its effectiveness in reliably differentiating between normal and abnormal cellular structures. These findings substantiate the promise of deep residual networks coupled with transfer learning as a powerful tool in advancing computer-aided diagnostic systems, thereby reinforcing early screening initiatives for cervical cancer.

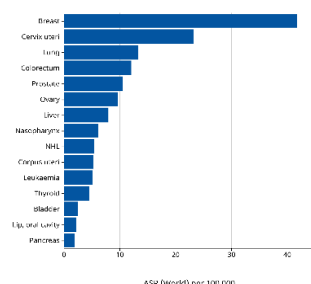


This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Pada data epidemiologi menunjukkan bahwa kanker serviks merupakan salah satu penyebab kematian tertinggi pada wanita di seluruh dunia. Pada negara berkembang, terutama di Indonesia, kanker serviks merupakan kanker terbanyak kedua pada wanita[1].

Age-Standardized Rate (World) per 100 000, Incidence, Both sexes, In 2022
Indonesia
(Top 15 cancer sites)



Cancer TODAY | IARC - <https://doi.org/10.1016/j.cancer.2023.10.011>
Data source: Cancer registry data, 2022
© 2023 Elsevier B.V. All rights reserved.



Gambar 1. Terdapat 15 jenis kanker paling umum pada wanita di Indonesia menurut WHO.

Berdasarkan Global Cancer Observatory (Globocan) dari World Health Organization (WHO) tahun 2022, kanker serviks adalah jenis kanker yang menduduki peringkat kedua tertinggi setelah kanker payudara yang paling banyak dialami oleh perempuan di Indonesia, dengan total 36.964 kasus baru atau sekitar 16,8% dari seluruh kasus kanker pada perempuan[2]. Sementara itu, menurut Kementerian Kesehatan Republik Indonesia, sekitar 70% kasus kanker serviks terdeteksi pada stadium lanjut, sehingga meningkatkan risiko kematian secara signifikan[3]. Kondisi ini menegaskan pentingnya deteksi dini.

Sebagai respons global, WHO melalui strategi eliminasi kanker serviks 2030 menargetkan pencapaian “90-70-90”: vaksinasi HPV pada 90 % anak perempuan di bawah usia 15 tahun, skrining berbasis metode berperforma tinggi pada 70 % perempuan usia 35 dan 45 tahun, serta pengobatan tuntas bagi 90 % perempuan yang terdiagnosis[4]. Di tingkat nasional, Kemenkes Indonesia menetapkan target skrining

75 % perempuan usia 30–69 tahun dengan metode DNA HPV dan memperluas cakupan melalui skrining berbasis fasilitas kesehatan primer maupun metode swab mandiri[5]. Pendeteksian dini melalui pemeriksaan Pap smear terbukti telah lama menjadi standar emas untuk skrining kanker serviks, namun interpretasi manual membutuhkan waktu yang lama dan juga rentan terhadap variabilitas diagnostik[6]. Untuk mengatasi hal tersebut, pengembangan sistem klasifikasi otomatis berbasis *deep learning* saat ini telah menjadi solusi yang menjanjikan[7]. Pendekatan *deep learning*, terkhusus pengolahan citra pada Convolutional Neural Network (CNN), telah menunjukkan kinerja superior dalam klasifikasi citra sel serviks karena kemampuannya dalam mengekstraksi fitur spasial citra tanpa pra-pemrosesan kompleks[8]. Salah satu strategi yang banyak digunakan dalam domain medis adalah *transfer learning*, yaitu dengan memanfaatkan model CNN yang telah dilatih sebelumnya (*pre-trained*) untuk tugas klasifikasi baru dengan jumlah data yang terbatas[9].

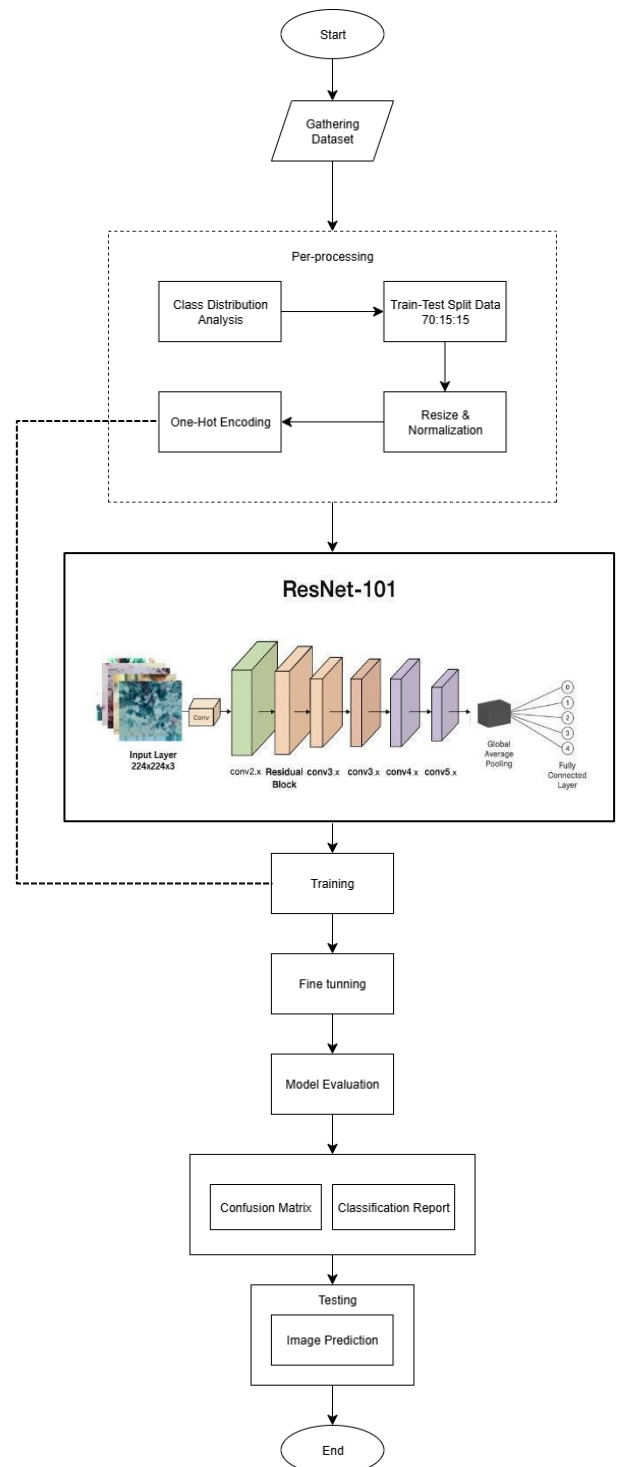
Beberapa studi terdahulu telah mengeksplorasi penggunaan arsitektur ResNet, termasuk ResNet50 yang telah terbukti efektif mengatasi *vanishing gradient* pada jaringan dalam dan mengekstraksi fitur visual kompleks[10]. Versi lanjutan, ResNet101, menawarkan kedalaman jaringan yang lebih tinggi untuk mengolah dan menangkap fitur deteksi pola visual yang lebih kompleks, dan telah digunakan dalam berbagai aplikasi klasifikasi dengan hasil akurasi yang menjanjikan[11]. Penerapan ResNet101 pada framework TensorFlow ini telah mampu menghasilkan performa yang kompetitif pada tugas klasifikasi multikelas citra medis[12]. Namun, pada kasus klasifikasi kanker serviks berbasis citra sel, khususnya Pap smear, akurasi model sering mengalami penurunan seiring bertambahnya jumlah kelas, sehingga diperlukan evaluasi mendalam terhadap performa model pada skenario multikelas [13]. Beberapa penelitian menyatakan bahwa akurasi tinggi (>95%) dapat dicapai dalam klasifikasi Pap smear lima kelas menggunakan pendekatan *hybrid* seperti, *multi-branch CNN*, atau penggabungan dari CNN dan transformer. Namun, metode tersebut memiliki kelemahan berupa kompleksitas arsitektur dan kebutuhan komputasi yang tinggi, sehingga kurang praktis untuk implementasi di lingkungan klinis[14].

Oleh karena itu, penelitian ini bertujuan untuk mengevaluasi kinerja arsitektur ResNet101 berbasis *transfer learning* dengan pendekatan *feature extraction* dalam tugas klasifikasi Pap smear lima kelas menggunakan dataset terbuka dari Kaggle. Penelitian ini diharapkan dapat memberikan kontribusi pada pengembangan sistem diagnosis berbantuan komputer (CAD) yang efisien, sederhana, dan presisi tinggi untuk deteksi kanker serviks secara dini.

II. METODE

Bagian metode akan memaparkan secara rinci setiap tahapan penelitian yang dilakukan, mulai dari persiapan data, pemilihan teknik, hingga prosedur analisis.

Selain itu, bagian ini juga akan menguraikan teori-teori yang menjadi dasar dari pendekatan yang digunakan dalam studi ini. Uraian tersebut disusun secara sistematis untuk memberikan gambaran yang jelas mengenai alur penelitian serta justifikasi dalam pemilihan metode.



Gambar 2. Alur penelitian mulai dari pra-pemrosesan, pelatihan pada model arsitektur ResNet 101, fine-tuning, evaluasi, hingga pengujian.

A. Dataset

Evaluasi dilakukan menggunakan dataset SIPaKMeD yang tersedia secara publik di Kaggle. Dataset pada penelitian ini merupakan pengembangan dari dataset sebelumnya yang hanya memiliki 4.049 gambar. Hingga saat ini, beberapa studi sebelumnya masih menggunakan dataset lama, sehingga belum ada penelitian yang memanfaatkan dataset yang telah dikembangkan ini[15]. Dataset hasil pengembangan tersebut melalui proses augmentasi hingga berjumlah 25.000 citra dengan pendekatan klasifikasi multi-class ke dalam lima kategori sel serviks: cervix_pab (Parabasal) dan cervix_sfi (Superficial-Intermediate) sebagai sel Normal, cervix_mep (Metaplastic) sebagai sel Jinak (*Benign*) dan cervix_dyf (Dyskeratotic), cervix_koc (Koilocytotic) sebagai sel Abnormal[16].

Penggunaan dataset augmentasi dipilih untuk mengatasi ketidakseimbangan data (*class imbalance*) serta dapat meningkatkan keragaman variasi visual yang mungkin dijumpai di kondisi klinis nyata, seperti rotasi sel, perbedaan pencahayaan, dan variasi bentuk inti sel[17]. Dengan augmentasi, risiko overfitting pada dataset kecil dapat ditekan, sementara robustness model terhadap variasi data dunia nyata dapat ditingkatkan. Pendekatan ini sejalan dengan praktik terbaik dalam pengembangan sistem diagnosis berbantuan komputer *Computer-Aided Diagnosis* (CAD) untuk deteksi dini kanker serviks, di mana keterbatasan jumlah citra dan tingginya variasi visual menjadi tantangan utama[18].

TABEL I
RINCIAN ISI DATASET

Dataset		
Sub-kelas	Total	Deskripsi
cervix_dyf (Dyskeratotic)	5.000	Pertumbuhan sel yang tidak normal
cervix_koc (Koilocytotic)	5.000	Sel yang menunjukkan perubahan akibat infeksi virus (misalnya HPV)
cervix_mep (Metaplastic)	5.000	Sel yang berubah dari satu jenis ke jenis lainnya (prakanker)
cervix_pab (Parabasal)	5.000	Sel skuamosa yang belum matang
cervix_sfi (Superficial-Intermediate)	5.000	Sel skuamosa yang lebih matang

B. Pra-Pemrosesan Data

Penelitian ini melakukan langkah-langkah prapemrosesan untuk meningkatkan kualitas dan konsistensi data sebelum diolah oleh model pembelajaran[19]. Beberapa metode prapemrosesan yang diterapkan dalam studi ini meliputi

analisis distribusi kelas (*class distribution analysis*), untuk memahami proporsi setiap kategori dalam dataset dan mengidentifikasi adanya ketidakseimbangan kelas, yang mana jika diabaikan dapat merusak performa klasifikasi dan mengarah pada bias prediksi[20]. Selanjutnya dilakukan pembagian data (*train-test split data*) yang dibagi ke dalam rasio 70 % pelatihan, 15 % validasi, dan 15 % pengujian untuk menjaga reliabilitas evaluasi model. Strategi serupa juga digunakan oleh studi “Privacy Preserved Cervical Cancer Detection Using CNNs on Pap Smear Images”, yang melaporkan akurasi uji sekitar 91% dengan pembagian 70 : 15 : 15 data[21]. Setiap citra kemudian diproses melalui tahapan penskalaan dan normalisasi (*resize & normalization*), dimana setiap gambar diubah ukurannya (misalnya ke 224×224 piksel) agar konsisten dengan input yang diharapkan oleh arsitektur *neural network*, dan nilai pikselnya dinormalisasi ke rentang [0, 1] untuk stabilitas numerik selama optimasi model[22]. Selanjutnya, dilakukan *One-Hot Encoding* terhadap label kategori, yang dapat mengubah label diskrit menjadi representasi vektor biner ini esensial untuk kebutuhan fungsi loss seperti categorical *cross-entropy* dalam setting klasifikasi multi-kelas[23]. Dengan demikian, rangkaian prapemrosesan ini memastikan data yang digunakan berada dalam format optimal, sehingga mampu meningkatkan performa, akurasi, dan generalisasi model pada tugas klasifikasi citra Pap smear.

TABEL II
PEMBAGIAN DATA TRAIN, TEST, VALIDATION

Description	Data Training	Data Testing	Data Validation	Total
Jumlah	70%	15%	15%	100%
Proporsi	17.500	3.750	3.750	25.000

C. Perancangan Arsitektur CNN

Menerapkan struktur lapisan pada arsitektur ResNet101, yang dirancang untuk klasifikasi gambar pada ImageNet, ke dalam penelitian ini[24]. Dengan memanfaatkan pendekatan *Transfer Learning* (TL) digunakan untuk memanfaatkan bobot pra-latih (*pretrained weights*) dari dataset besar seperti ImageNet, sehingga mempercepat konvergensi dan mengurangi kebutuhan data pelatihan yang besar[25]. Selanjutnya, dilakukan *Feature Extraction* (FE) dengan menggunakan lapisan-lapisan konvolusional ResNet-101 sebagai “*fixed feature extractor*”; hanya lapisan klasifikasi akhir (*fully connected*) yang diadaptasi (*fine-tuning*) agar sesuai dengan jumlah kelas target pada dataset ini[26].

1. *Lapisan Input*: Citra masukan memiliki dimensi tetap 224×224 piksel dan tiga saluran warna (RGB), disesuaikan dengan format arsitektur ResNet-101 pra-latih yang digunakan[27].
2. *Konvolusi Awal & Pooling*: Dilakukan konvolusi awal dengan kernel 7×7 dan stride 2 untuk

mengekstraksi fitur fundamental, diikuti *max pooling* 3×3 (stride 2) guna mereduksi dimensi spasial sambil mempertahankan informasi penting[28].

3. *Struktur Blok Residual (Bottleneck)*: ResNet-101 terdiri atas sekitar 3 blok *bottleneck*, terbagi ke dalam empat tahap utama. Setiap blok meliputi tiga lapisan yaitu 1×1 *convolution* (reduction) yang mengecilkan dimensi kanal $\rightarrow 3 \times 3$ *convolution* untuk mengekstraksi fitur spesial $\rightarrow 1 \times 1$ *convolution* (expansion) untuk mengembalikan dimensi awal, dengan *skip connection* menambah input langsung ke output blok untuk menjaga stabilitas gradien dan memperlancar propagasi fitur[29].

D. Pelatihan Model

Pada tahap pelatihan model, penelitian ini menggunakan ResNet-101 *pretrained* ImageNet sebagai *backbone*, dengan konvolusi awal dibekukan sehingga hanya lapisan klasifikasi tambahan yang dilatih. Kepala klasifikasi tersusun dari lapisan Flatten $\rightarrow$ Dense (256 dan 64 neuron, ReLU) $\rightarrow$ Dropout (0.25) $\rightarrow$ Dense output (5 neuron, Softmax). Model dikompilasi dengan *optimizer Adamax* ($\text{lr}=0.002$), fungsi *loss categorical_crossentropy*, dan metrik akurasi. Pelatihan berjalan selama 50 epoch menggunakan *train_gen* dan *valid_gen*, dengan *EarlyStopping* (*patience*=5) untuk mencegah *overfitting* dan *ModelCheckpoint* untuk menyimpan bobot terbaik yang didapatkan pada pelatihan model.

Pendekatan serupa juga diterapkan oleh Adhikari dkk.[30] yang memanfaatkan *transfer learning* dengan MobileNetV2 dan menerapkan *callback* untuk pemilihan model terbaik, sehingga mampu mencapai akurasi 93,65% pada klasifikasi kanker serviks. Hal ini menunjukkan bahwa strategi *transfer learning* dengan kombinasi *callback* efektif dalam meningkatkan performa model klasifikasi citra medis.

E. Fine-Tuning

Pada tahap ini, metode penyempurnaan (*fine-tuning*) dilakukan dengan cara membekukan sebagian besar lapisan pada arsitektur ResNet-101 dan hanya melatih ulang beberapa lapisan terakhir. Strategi ini dimaksudkan untuk mempertahankan representasi fitur umum yang telah dipelajari dari dataset pra-latih ImageNet, sekaligus menyesuaikannya dengan karakteristik khusus pada dataset penelitian. Penggunaan *learning rate* kecil ($1e-5$) bertujuan untuk menghindari perubahan drastis pada bobot awal sehingga mencegah terjadinya *catastrophic forgetting*. Selain itu, integrasi *callback* berupa *EarlyStopping* dan *ModelCheckpoint* memastikan proses pelatihan berhenti ketika model tidak lagi mengalami peningkatan, sekaligus menyimpan bobot terbaik yang diperoleh. Dengan pendekatan ini, proses *fine-tuning* dapat meningkatkan kemampuan generalisasi model terhadap data baru, menjaga efisiensi komputasi, serta menghasilkan model dengan performa optimal pada data validasi.

F. Evaluasi Model

Evaluasi performa model dilakukan menggunakan data uji yang sebelumnya telah dipisahkan untuk mengukur seberapa baik model mengklasifikasikan gambar Pap Smear ke dalam lima kategori kelas, yaitu *cervix_dy*, *cervix_koc*, *cervix_mep*, *cervix_pab*, *cervix_sfi*. Metrik evaluasi yang digunakan meliputi akurasi, presisi, recall, dan F1-score. Akurasi digunakan untuk mengukur persentase prediksi yang benar dari keseluruhan data uji. Presisi mengukur proporsi prediksi yang benar dari seluruh prediksi positif yang dihasilkan model, sedangkan recall mengukur seberapa banyak data positif yang berhasil teridentifikasi oleh model. F1-score digunakan sebagai ukuran keseimbangan antara presisi dan recall, terutama penting pada kondisi distribusi kelas yang tidak seimbang. Selain itu, confusion matrix digunakan untuk mengevaluasi hasil klasifikasi secara lebih detail berdasarkan jumlah prediksi benar dan salah pada tiap kelas.

$$\text{Accuracy\%} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision\%} = \frac{TP}{TP + FP}$$

$$\text{Recall\%} = \frac{TP}{TP + FN}$$

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

Penjelasan:

- TP (True Positive) = kasus positif yang benar terdeteksi positif.
- TN (True Negative) = kasus negatif yang benar terdeteksi negatif.
- FP (False Positive) = kasus negatif yang salah terdeteksi positif.
- FN (False Negative) = kasus positif yang salah terdeteksi negatif.

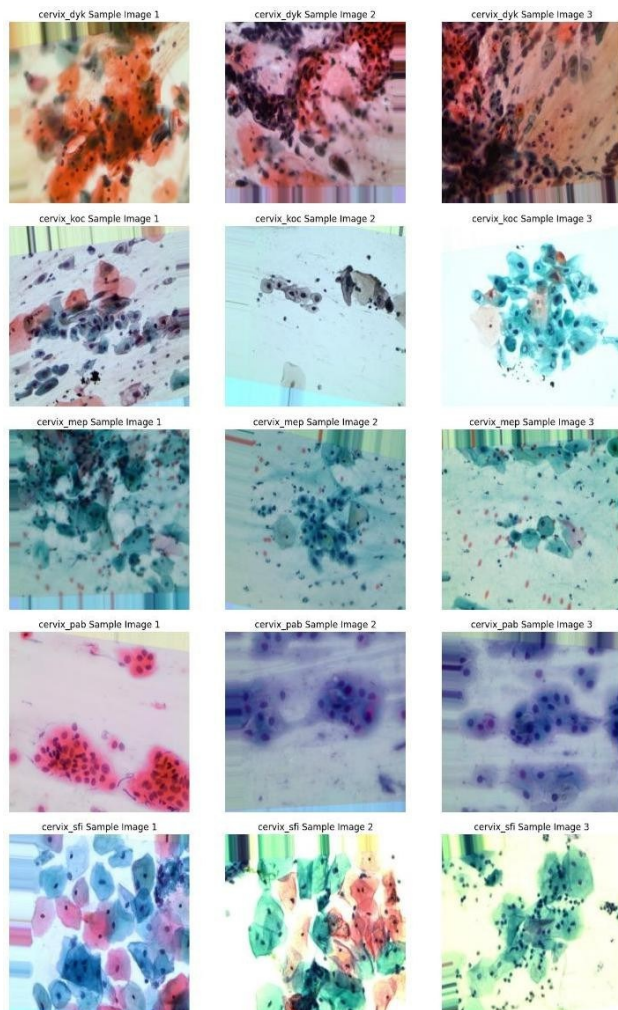
G. Analisis Hasil

Hasil evaluasi model CNN dengan arsitektur ResNet101 berbasis transfer learning menunjukkan performa klasifikasi yang sangat tinggi terhadap seluruh kelas dalam dataset. Berdasarkan hasil pengujian, model mencapai akurasi keseluruhan 99,7%, dengan nilai precision, recall, dan f1-score hampir sempurna di semua kelas. Seluruh kelas memperoleh precision dan recall sebesar 1,00, kecuali pada kelas 1 yaitu *cervix_koc* (Koilocytotic) dengan recall sebesar 0,99 namun tetap menghasilkan precision dan f1-score sebesar 1,00. Nilai macro average dan weighted average juga mencapai 1,00, menandakan kinerja model konsisten dan seimbang tanpa bias antar kelas.

III. HASIL DAN PEMBAHASAN

A. Dataset

Hasil pelatihan menunjukkan bahwa model mampu memanfaatkan keragaman visual dalam Multi Cancer Dataset secara efektif. Distribusi yang seimbang pada lima kategori sel Dyskeratotic, Koilocytotic, Metaplastic, Parabasal, dan Superficial-Intermediate berkontribusi terhadap konsistensi akurasi tinggi pada seluruh kelas, tanpa kecenderungan bias.

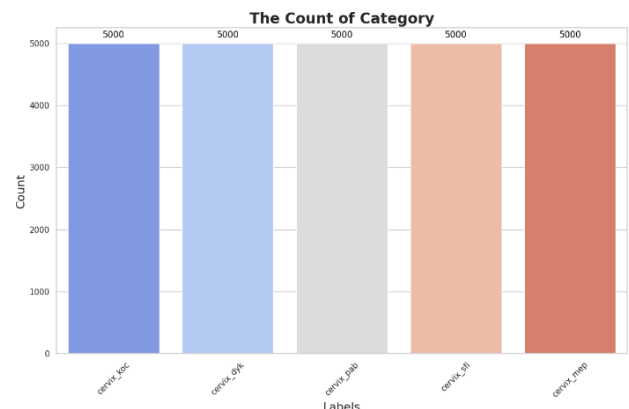


Gambar 3. Gambar ini adalah visualisasi gambar pada dataset Cervical Cancer pada dataset Multi Cancer berupa Pap Smear Images.

B. Pra-pemrosesan Data

Sebelum pelatihan model dilakukan, tahap prapemrosesan data perlu dilakukan dengan tujuan untuk meningkatkan kualitas data agar sesuai dengan kebutuhan model deep learning. Beberapa tahapan yang umumnya dilakukan antara lain:

1. *Class Distribution Analysis*: dilakukan untuk mengevaluasi jumlah data pada masing-masing kategori sehingga potensi ketidakseimbangan kelas dapat terdeteksi sejak awal. Jika masalah *imbalance* diabaikan, model cenderung memberikan prediksi bias pada kelas mayoritas dan menurunkan generalisasi klasifikasi. Hasil pemeriksaan menunjukkan bahwa dataset yang digunakan bersifat seimbang (*balanced*) sehingga dapat mendukung proses pelatihan model tanpa perlu penanganan tambahan seperti teknik *oversampling* atau *undersampling*.



Gambar 4. Keberagaman jumlah data tiap class.

2. *Train-Test Split Data*: dilakukan dengan rasio 70:15:15 untuk pengujian. Pembagian ini memastikan bahwa model dilatih pada data yang cukup besar, sementara validasi dan pengujian tetap memiliki proporsi representatif guna menilai performa model secara objektif. Pendekatan serupa telah digunakan dalam studi "Privacy Preserved Cervical Cancer Detection Using CNNs on Pap Smear Images", yang melaporkan akurasi uji hingga 91%[26].

TABEL III
PEMBAGIAN DATA TRAIN, TEST, VALIDATION

Informasi	Data Training	Data Testing	Data Validation	Total
Jumlah	70%	15%	15%	100%
Proporsi	17.500	3.750	3.750	25.000
/cervix_dyk	3.500	750	750	5.000
/cervix_koc	3.500	750	750	5.000
/cervix_mep	3.500	750	750	5.000
/cervix_pab	3.500	750	750	5.000
/cervix_sfi	3.500	750	750	5.000

3. *Resize & Normalization*: diterapkan dengan tujuan seluruh gambar memiliki ukuran seragam, misalnya 224×224 piksel, sehingga sesuai dengan arsitektur

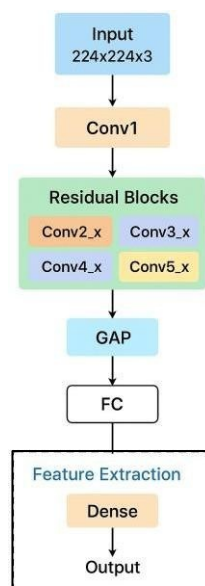
input dari CNN yang digunakan. Selain itu, nilai piksel setiap citra dinormalisasi ke rentang $[0, 1]$ untuk mempercepat konvergensi model, mengurangi risiko *exploding/vanishing gradient*, serta menjaga stabilitas numerik selama proses optimasi.

4. *On-Hot Encoding*: digunakan untuk mengonversi kategori diskrit menjadi representasi vektor biner. Tahap ini sangat penting dalam setting klasifikasi multi-kelas, karena memungkinkan fungsi *loss* seperti *categorical cross-entropy* menghitung error secara tepat antara prediksi probabilistik dan label ground truth.

Dengan demikian, pada proses prapemrosesan menyiapkan data dalam kondisi optimal sehingga model bekerja lebih efisien, meningkatkan akurasi prediksi, dan memperkuat generalisasi terhadap variasi citra Pap smear.

C. ResNet101

Model yang digunakan pada penelitian ini adalah ResNet101 yang berfungsi sebagai *feature extractor* dengan bobot awal (*pretrained weights*) dari ImageNet. ResNet101 dipilih karena memiliki arsitektur yang dalam dengan kemampuan mengekstraksi fitur citra secara hierarkis dan telah terbukti unggul pada berbagai tugas klasifikasi citra.



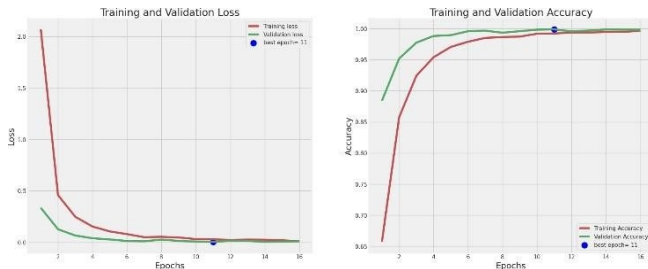
Gambar 5. Tahapan pemrosesan arsitektur ResNet101 per-lapisan dalam Feature Extraction dan Transfer Learning.

Arsitektur ResNet101 terdiri dari 101 lapisan yang dibangun berdasarkan konsep residual learning, yang dirancang khusus untuk mengatasi permasalahan vanishing gradient pada jaringan yang sangat dalam. Proses pengolahan data pada model ini dilakukan melalui beberapa tahap bertahap, yang akan dijelaskan secara lebih rinci berikut, sebelum data mencapai tahap akhir prediksi, yaitu:

1. *Input Layer (224x224x3)*: gambar Pap smear berukuran 224×224 piksel dengan tiga kanal warna (RGB) dimasukkan ke jaringan. Tahap ini mencocokkan format data dengan arsitektur ResNet101 yang sudah *pre-trained* pada ImageNet.
2. *Conv1 (7×7 Convolution + BatchNorm + ReLU + MaxPooling)*: Lapisan awal menggunakan kernel 7×7 dengan *stride* 2 untuk mengekstraksi fitur dasar seperti tepi (*edges*) dan tekstur kasar. Dilanjutkan dengan *batch normalization* untuk stabilisasi distribusi aktivasi, *ReLU* sebagai fungsi aktivasi non-linear, dan *max pooling* untuk mengurangi resolusi spasial.
3. *Residual Blocks (Conv2_x, Conv3_x, Conv4_x, Conv5_x)*: terdiri dari 33 blok residual berbasis bottleneck, di mana setiap blok memiliki tiga lapisan konvolusi berurutan (1×1, 3×3, 1×1). Pada tahap Conv2_x, jaringan mengekstraksi pola dasar seperti garis dan bentuk sederhana. Selanjutnya, Conv3_x menangkap pola menengah, misalnya tepi berulang atau struktur sel yang lebih kompleks. Tahap Conv4_x berfokus pada identifikasi pola yang lebih rumit, termasuk distribusi warna dan tekstur khas pada sel serviks. Pada tahap akhir, Conv5_x mengenali pola yang sangat spesifik yang menjadi pembeda antar jenis sel, seperti variasi bentuk inti dan sitoplasma. Seluruh proses ini didukung oleh mekanisme *skip connection*, yang memungkinkan gradien mengalir langsung ke lapisan awal sehingga stabilitas pembelajaran tetap terjaga meskipun kedalaman jaringan sangat tinggi.
4. *Global Average Pooling (GAP)*: berfungsi untuk menghitung rata-rata nilai aktivasi pada setiap feature map, menghasilkan representasi vektor fitur berdimensi rendah yang kaya informasi, sekaligus mengurangi risiko *overfitting*.
5. *Fully Connected layer (FC)*: yaitu lapisan *dense tradisional* yang menghubungkan semua neuron dari layer sebelumnya ke semua neuron di layer berikutnya, sehingga memungkinkan pembelajaran representasi yang kompleks dan menyeluruh dari data.
6. *Feature Extraction*: Pada tahap ini, bobot seluruh lapisan konvolusi ResNet101 dibekukan sehingga hanya berfungsi mengekstraksi fitur, yang kemudian diringkas melalui Global Average Pooling (GAP). Pada tahap Transfer Learning, lapisan fully connected bawaan diganti dengan arsitektur kustom: Flatten untuk mengubah fitur GAP menjadi 1D, Dense(256) + Dropout untuk penggabungan fitur sekaligus mencegah *overfitting*, Dense(64) + Dropout untuk reduksi dimensi, dan Dense(5, Softmax) sebagai keluaran untuk klasifikasi lima tipe sel serviks.

D. Training Model

Proses pelatihan dilakukan selama 50 epoch untuk mengoptimalkan pembaruan bobot pada lapisan yang dilatih, memastikan model mampu mempelajari pola fitur secara menyeluruh, sekaligus memantau performa guna mencegah *overfitting* melalui mekanisme validasi.



Gambar 6. Grafik Perbandingan Loss dan Akurasi pada Data Latih dan Validasi selama Proses Pelatihan Model

Grafik pada Gambar menunjukkan perubahan nilai *loss* dan akurasi pada data latih serta validasi selama proses pelatihan model. Kurva *loss* memperlihatkan penurunan tajam pada epoch awal, yang mengindikasikan bahwa model dengan cepat mempelajari pola dasar dari data. Setelah itu, *loss* cenderung stabil mendekati nol, baik pada data latih maupun validasi. Sementara itu, kurva akurasi menunjukkan peningkatan signifikan sejak epoch awal hingga mendekati 100%, dengan titik optimal dicapai pada epoch ke-16. Pola ini menandakan bahwa model tidak mengalami *overfitting* secara signifikan dan mampu mempertahankan kinerja tinggi pada data yang tidak dilatih, sejalan dengan hasil evaluasi *confusion matrix* dan *classification report* yang mencapai akurasi 99,7%.

Proses pelatihan dengan *training epoch* sebanyak 50 kali dilakukan dalam dua kali percobaan sebelum tahap *fine-tuning*. Pada percobaan pertama, titik optimal tercapai di epoch ke-26, pada percobaan kedua performa terbaik sudah diperoleh di epoch ke-16. Hal ini menunjukkan terjadinya percepatan konvergensi dan peningkatan efisiensi pelatihan pada percobaan berikutnya.

E. Fine-Tuning

Model arsitektur ResNet101 yang telah dilatih sebelumnya dan disimpan dalam format (.keras) kemudian menjalani proses *fine-tuning* untuk meningkatkan performa secara lebih spesifik dan teoptimasi. Proses ini dilakukan dengan menyesuaikan parameter pelatihan secara lebih sensitif agar model dapat beradaptasi secara optimal terhadap karakteristik unik dari dataset yang digunakan.

TABEL IV
FINE-TUNING

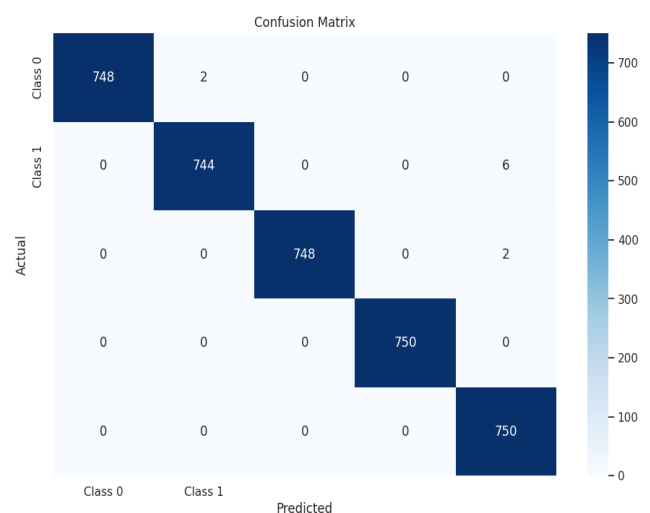
Name	Parameter Value
Fine-Tuning	0,5
Optimizer	Adamax
Learning Rate	0,00001
Epoch	50

Memanfaatkan mekanisme *callbacks* selama proses *fine-tuning* memungkinkan model dapat beradaptasi secara lebih efisien dan cepat terhadap data validasi dan menghindari masalah seperti *overfitting*. Dalam penelitian ini, digunakan pula dua jenis *callbacks*, yaitu EarlyStopping dengan parameter *patience* sebesar 5, pemantauan pada *validation loss*, serta opsi *restore best weights* untuk memastikan bobot terbaik tetap digunakan, dan ModelCheckpoint untuk menyimpan model dengan kinerja validasi terbaik secara otomatis. Kombinasi strategi ini memastikan proses pelatihan berhenti pada saat yang optimal serta mempertahankan performa model meskipun terjadi sedikit penurunan, sehingga pada tahap akhir pelatihan *fine-tuning* diperoleh akurasi sebesar 99,6%.

F. Model Evaluation

Model ResNet101 tanpa *fine-tuning* dipilih sebagai model evaluasi karena menunjukkan akurasi yang sedikit lebih tinggi dengan pendekatan *Feature Extraction* (99,7%), di mana hanya lapisan atas (*fully connected/classifier*) yang dilatih. Sebaliknya, model dengan *fine-tuning* penuh menghasilkan akurasi sedikit lebih rendah (99,6%) meskipun membuka sebagian atau seluruh lapisan untuk dilatih ulang. Model tanpa *fine-tuning* terbukti memberikan kinerja keseluruhan yang lebih unggul, dengan waktu pelatihan lebih singkat, generalisasi lebih baik, serta minim risiko *overfitting* dibandingkan dengan model yang dilatih ulang.

1. *Confusion Matrix*: Hasil Confusion Matrix menunjukkan bahwa pelatihan model ResNet101 pre-trained dengan pendekatan Feature Extraction (pada lapisan fully connected/classifier) dalam transfer learning menghasilkan performa klasifikasi yang baik.



Gambar 7. Confusion Matrix Model ResNet101 dengan pendekatan *Feature Extraction* (*fully connected/classifier*).

Berdasarkan Gambar 7, ResNet101 dengan optimasi *feature extraction* dan *partial fine-tuning* mencapai akurasi tinggi serta kinerja seimbang pada kelima kelas. Hanya sedikit kesalahan klasifikasi, menunjukkan kemampuan model dalam merepresentasikan pola visual penting pada citra Pap Smear.

2. *Classification Report*: Hasil ini diperoleh dari model ResNet101 tanpa penerapan *fine-tuning* penuh, yang telah dilatih secara optimal dan menghasilkan metrik evaluasi berupa akurasi, recall, precision, serta F1-score.

TABEL V
CLASSIFICATION REPORT RESNET101

Classes	Precision	Recall	F1-Score
0	1.00	1.00	1.00
1	1.00	0.99	0.99
2	1.00	1.00	1.00
3	1.00	1.00	1.00
4	0.99	1.00	0.99

Berdasarkan Tabel V, hasil classification report menunjukkan performa model yang sangat optimal. Kelas 0, 2, dan 3 memperoleh precision, recall, dan F1-score sebesar 1.00, menandakan klasifikasi sangat baik. Pada kelas 1, recall sedikit menurun menjadi 0.99 dengan F1-score 0.99, sedangkan pada kelas 4 precision turun tipis ke 0.99 dengan recall tetap 1.00. Secara keseluruhan, model menunjukkan kapabilitas klasifikasi yang sangat tinggi dan konsisten pada mayoritas kelas. Selanjutnya, analisis lebih rinci terkait kesalahan prediksi *false negative* yang akan disajikan melalui tabel kesalahan.

TABEL VI
FALSE NEGATIVE PREDICTION

Class	Class Name	Prediction
0	Dyskeratotic	2
1	Koliocytotic	6
2	Metaplastic	2

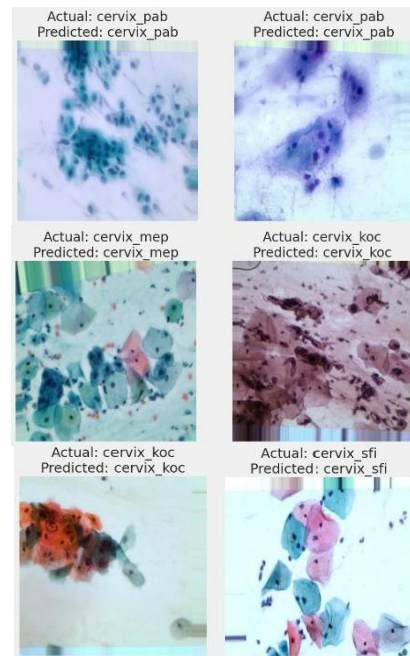
Analisis kesalahan pada Tabel VI menunjukkan adanya *false negative* pada beberapa kelas, yaitu 2 pada Dyskeratotic, 6 pada Koliocytotic, dan 2 pada Metaplastic. Jumlah ini relatif kecil dibandingkan total 750 citra per kelas, sehingga tidak berpengaruh signifikan terhadap akurasi keseluruhan sebesar 99,7%. Kesalahan terutama disebabkan oleh kemiripan morfologi antar sel abnormal yang membuat model sulit membedakan kelas tertentu, khususnya Koliocytotic dan Metaplastic, sebagaimana juga dilaporkan pada penelitian sebelumnya.

G. Image Prediction

1. Uji coba data dengan menggunakan metode acak dilakukan untuk mengevaluasi performa model secara lebih objektif. Pendekatan ini bertujuan untuk

memastikan bahwa model tidak hanya mampu akan mengenali pola dari data tertentu, tetapi juga dapat memprediksi data baru yang belum pernah dilihat sebelumnya.

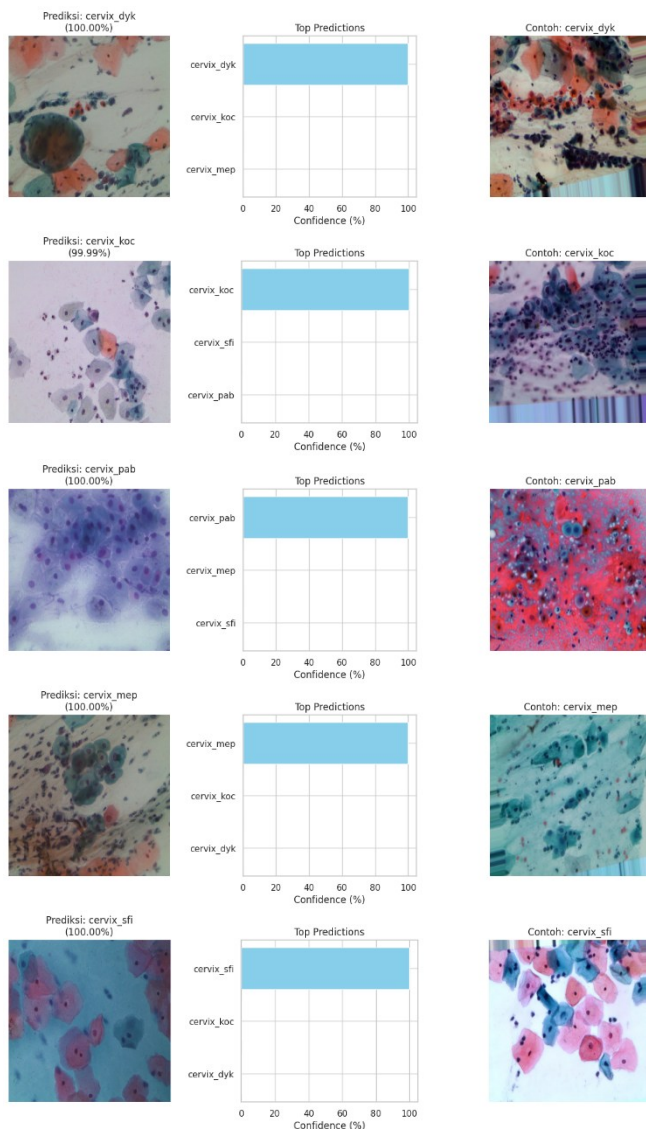
Dengan demikian, hasil akurasi yang diperoleh mencerminkan kemampuan generalisasi model terhadap keseluruhan dataset.



Gambar 8. Evaluasi performa model melalui uji coba data acak.

Berdasarkan Gambar 8, hasil uji coba menunjukkan bahwa model mampu melakukan prediksi dengan baik pada data uji yang dipilih secara acak. Keberhasilan prediksi pada data acak ini juga menjadi indikator bahwa model memiliki tingkat generalisasi yang baik, sehingga tidak hanya menghafal data latih, tetapi juga mampu menyesuaikan diri terhadap variasi data yang berbeda.

2. Pengujian data kedua dilakukan dengan menggunakan pendekatan upload data independen yang terdiri dari lima kelas. Pendekatan ini bertujuan untuk menilai kemampuan model dalam melakukan prediksi terhadap data baru yang sepenuhnya berada di luar distribusi data latih. Dengan kata lain, model diuji tidak hanya pada variasi internal dataset yang telah digunakan selama pelatihan, tetapi juga pada data eksternal yang lebih beragam, yang mungkin mengandung pola atau karakteristik yang sebelumnya belum pernah ditemui oleh model. Proses pengujian ini dilakukan secara langsung melalui running dan upload file menggunakan platform Google Colab.



Gambar 9. Evaluasi performa model melalui uji coba data independen.

Terbukti bahwa dengan melakukan dua metode uji coba, yaitu pengujian menggunakan data acak dan pengujian dengan data independen, model mampu menghasilkan tingkat akurasi yang konsisten dan dapat diandalkan. Hasil ini menunjukkan bahwa performa model tidak hanya baik pada data yang berasal dari dataset awal, tetapi juga tetap stabil ketika dihadapkan pada data baru yang bersifat independen. Dengan demikian, kedua metode uji coba

tersebut memperkuat validitas model dalam hal generalisasi dan memberikan keyakinan bahwa model siap diterapkan pada kondisi nyata.

IV. KESIMPULAN

Berdasarkan serangkaian pengujian yang telah dilakukan pada data Pap Smear, penelitian ini berhasil menunjukkan penggunaan arsitektur ResNet101 dengan pendekatan penggabungan *transfer learning* dan *feature extraction* melalui lapisan *fully connected*, serta penerapan mekanisme *callback* mampu menghasilkan performa klasifikasi yang tinggi, yakni 99,7%. Strategi ini terbukti lebih optimal ketika tidak dilakukan *fine-tuning* penuh, sehingga model tetap mampu mempertahankan akurasi tinggi tanpa kehilangan kemampuan generalisasi.

Selain itu, pendekatan serupa juga diterapkan pada MobileNetV2 dengan mekanisme *callback* untuk pemilihan model terbaik selama proses pelatihan. Teknik ini memungkinkan diperolehnya performa yang stabil dengan tingkat akurasi sebesar 93,65%. Hasil ini menegaskan bahwa dengan memanfaatkan *pre-trained* model yang dikombinasikan dengan strategi pelatihan yang tepat mampu meningkatkan performa tanpa memerlukan pelatihan ulang seluruh jaringan dari awal.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa pemilihan arsitektur *backbone* yang kuat, dikombinasikan dengan teknik *transfer learning*, sangat efektif dalam meningkatkan akurasi klasifikasi citra Pap Smear. Hal ini menunjukkan potensi besar penggunaan deep learning dalam membantu deteksi dini kanker serviks secara lebih cepat, akurat, dan andal.

DAFTAR PUSTAKA

- [1] S. Zhang, H. Xu, L. Zhang, and Y. Qiao, "Cervical cancer: Epidemiology, risk factors and screening," *Chinese Journal of Cancer Research*, vol. 32, no. 6, pp. 720–728, 2020, doi: 10.21147/j.issn.1000-9604.2020.06.05.
- [2] World Health Organization, *Cancer Today: Data Visualization Tools for Exploring the Global Cancer Burden in 2022*. International Agency for Research on Cancer, 2022. [Online]. Available: <https://gco.iarc.fr/today>
- [3] Kementerian Kesehatan Republik Indonesia, "Kemenkes Tegaskan Komitmen Eliminasi Kanker Serviks: 36 Ribu Kasus Baru Terdeteksi Setiap Tahun," *Kementerian Kesehatan RI*, Feb. 2024. [Online]. Available: <https://www.kemkes.go.id>
- [4] World Health Organization, *Global Strategy to Accelerate the Elimination of Cervical Cancer as a Public Health Problem*. Geneva: WHO, 2020. [Online]. Available: <https://www.who.int/publications/i/item/978924001>

- [5] PressRelease.id, "Skining kanker serviks modal utama turunkan angka kematian perempuan," *PressRelease.id*, Feb. 2024. [Online]. Available: <https://pressrelease.kontan.co.id/news/skrining-kanker-serviks-modal-utama-tanggulangi-kematian-perempuan>
- [6] S. L. Tan, G. Selvachandran, W. Ding, R. Paramesran, and K. Kotecha, "Cervical Cancer Classification From Pap Smear Images Using Deep Convolutional Neural Network Models," *Interdiscip Sci*, vol. 16, no. 1, pp. 16–38, Mar. 2024, doi: 10.1007/s12539-023-00589-5.
- [7] N. Al Mudawi and A. Alazeb, "A Model for Predicting Cervical Cancer Using Machine Learning Algorithms," *Sensors*, vol. 22, no. 11, Jun. 2022, doi: 10.3390/s22114132.
- [8] P. Jiang *et al.*, "A systematic review of deep learning-based cervical cytology screening: from cell identification to whole slide image analysis," *Artif Intell Rev*, vol. 56, pp. 2687–2758, Nov. 2023, doi: 10.1007/s10462-023-10588-z.
- [9] S. Dayalane, S. Murugesan, S. K. Mathivanan, H. Rajadurai, K. Panneer Selvam, and M. A. Shah, "Cervical Cancer Classification Using Deep Learning Approach Using Colposcopy Images," *Neural Process Lett*, vol. 57, no. 4, Aug. 2025, doi: 10.1007/s11063-025-11770-w.
- [10] L. Liu, J. Liu, Q. Su, Y. Chu, H. Xia, and R. Xu, "Performance of artificial intelligence for diagnosing cervical intraepithelial neoplasia and cervical cancer: a systematic review and meta-analysis," *EClinicalMedicine*, vol. 80, Feb. 2025, doi: 10.1016/j.eclim.2024.102992.
- [11] B. Z. Wubineh, A. Rusiecki, and K. Halawa, "Classification of cervical cells from the Pap smear image using the RES_DCGAN data augmentation and ResNet50V2 with self-attention architecture," *Neural Comput Appl*, Dec. 2024, doi: 10.1007/s00521-024-10404-x.
- [12] Dr. Madhur Jain, Mayank Singh Bora, Sameer Chandnani, Sanidhay Grover, and Shivank Sadwal, "Comparison of VGG-16, VGG-19, and ResNet- 101 CNN Models for the purpose of Suspicious Activity Detection," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, pp. 121–130, Jan. 2023, doi: 10.32628/cseit2390124.
- [13] S. D. Roy, A. Mohanta, T. Edema, N. Nath, and M. K. Bhowmik, "Classification of Multiclass Pap Smear Image and Asymmetric Characterization for Prediction of Severity of Cervical Abnormality," *Int J Biomed Imaging*, vol. 2025, no. 1, 2025, doi: 10.1155/ijbi/9860677.
- [14] M. Al, M. Hasan, J. Shin, T. Baba, A. Saleh, and M. Miah, "Cervical Cancer Detection Using Multi- Branch Deep Learning Model", doi: 10.48550/arXiv.2408.10498.
- [15] M. M. Rahaman *et al.*, "DeepCervix: A Deep Learning-based Framework for the Classification of Cervical Cells Using Hybrid Deep Feature Fusion Techniques," Aug. 2021, doi: 10.1016/j.compbiomed.2021.104649.
- [16] D. Desiani, D. Fitriah, and A. A. Supianto, "Multiclass Classification of Cervical Pap Smear Images Using Deep Learning-Based Model," *Traitement du Signal*, vol. 40, no. 2, pp. 457–463, Apr. 2023, doi: 10.18280/ts.400204.
- [17] O. Attallah, "CerCan-Net: Cervical cancer classification model via multi-layer feature ensembles of lightweight CNNs and transfer learning," *Expert Syst Appl*, vol. 229, Nov. 2023, doi: 10.1016/j.eswa.2023.120624.
- [18] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, Jun. 2022, doi: 10.1016/j.gltp.2022.04.020.
- I. A. Reshma *et al.*, "Finding a Suitable Class Distribution for Building Histological Images Datasets Used in Deep Model Training—The Case of Cancer Detection," *J Digit Imaging*, vol. 35, no. 5, pp. 1326–1349, Oct. 2022, doi: 10.1007/s10278-022-00618-7.
- S. Alsubai *et al.*, "Privacy Preserved Cervical Cancer Detection Using Convolutional Neural Networks Applied to Pap Smear Images," *Comput Math Methods Med*, vol. 2023, no. 1, Jan. 2023, doi: 10.1155/2023/9676206.
- W. Liu *et al.*, "Is the aspect ratio of cells important in deep learning? A robust comparison of deep learning methods for multi-scale cytopathology cell image classification: from convolutional neural networks to visual transformers," Nov. 2021, [Online]. Available: <http://arxiv.org/abs/2105.07402>
- S. Dhawan, K. Singh, and M. Arora, "Cervix image classification for prognosis of cervical cancer using deep neural network with transfer learning," *EAI Endorsed Trans Pervasive Health Technol*, vol. 7, no. 27, 2021, doi: 10.4108/eai.12-4-2021.169183.
- Z. Niswati, R. Hardatin, M. N. Muslimah, and S. N. Hasanah, "Perbandingan Arsitektur ResNet50 dan ResNet101 dalam Klasifikasi Kanker Serviks pada Citra Pap Smear," *Faktor Exacta*, vol. 14, no. 3, p. 160, Oct. 2021, doi: 10.30998/faktorexacta.v14i3.10010.
- H. E. Kim, A. Cosa-Linan, N. Santhanam, M. Jannesari, M. E. Maros, and T. Ganslandt, "Transfer learning for medical image classification: a literature review," Dec. 01, 2022, *BioMed Central Ltd*, doi: 10.1186/s12880-022-00793-7.
- A. Davila, J. Colan, and Y. Hasegawa, "Comparison of fine-tuning strategies for transfer learning in medical image classification," Jun. 2024, doi: 10.1016/j.imavis.2024.105012.
- H. Alquran *et al.*, "Cervical Cancer Classification Using Combined Machine Learning and Deep Learning Approach," *Computers, Materials and Continua*, vol. 72, no. 3, pp. 5117–5134, 2022, doi: 10.32604/cmc.2022.025692.
- M. Shafiq and Z. Gu, "Deep Residual Learning for Image Recognition: A Survey," Sep. 01, 2022, MDPI, doi: 10.3390/app12188972.
- C. Rutili de Lima, S. G. Khan, S. H. Shah, and L. Ferri, "Mask region-based CNNs for cervical cancer progression diagnosis on pap smear examinations," *Heliyon*, vol. 9, no. 11, Nov. 2023, doi: 10.1016/j.heliyon.2023.e21388.
- H. Göker, "Detection Of Cervical Cancer From Uterine Cervix Images Using Transfer Learning Architectures," *Eskişehir Technical University Journal of Science and Technology A - Applied Sciences and Engineering*, vol. 25, no. 2, pp. 222–239, Jun. 2024, doi:10.18038/estubtda.1384489.
- S. Alsubai *et al.*, "Privacy Preserved Cervical Cancer Detection Using Convolutional Neural Networks Applied to Pap Smear Images," *Comput Math Methods Med*, vol. 2023, no. 1, Jan. 2023, doi: 10.1155/2023/9676206