iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81582

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Automated Cervical Cancer Detection Using Deep Learning: An Efficient Net-Based Framework with Transfer Learning on Pap Smear Images

Ch. Madhuri, Dr.U.Sathish Kumar, N. Sateesh Babu, B.Naga Veera, K. Sidhu
Department of Cyber Security Acharya Nagarjuna University Guntur, Andhra Pradesh, India

Abstract: ***Cervical cancer persists as one of the foremost causes of cancer-related mortality among women globally, with over 600,000 incident cases reported annually. Conventional Papanicolaou (Pap) smear-based screening demandsextensivepathologistexpertise, isinherentlytime- consuming, and is susceptible to inter-observer inconsistency. Thisinvestigationpresentsasystematicdeep learning framework for the automated classification of cervical cell images into five morphologically distinct categories—Dyskeratotic, Koilocytotic, Metaplastic, Parabasal, and Superficial-Intermediate—using the EfficientNet-B0 convolutional neural network in conjunction with a two-stage transfer learning strategy. Experiments conducted on the publicly available SIPaKMeD benchmark demonstrate that the proposed systemattains95.0% classification accuracywitha macro- averagedF1-scoreof94.4%,surpassingseveralestablished single-modelbaselines. Anadvancedaugmentation pipeline specifically designed to simulate laboratory staining variability, slide-preparation artefacts, and microscope orientation differences contributes materially to the observed generalisation capability. The trained model is deployed within a Flask-based web application that furnishes real-time per-class confidence scores alongside bar-chart visualisation, facilitating rapid clinical interpretation. The study further enumerates current limitations and delineates a concrete research trajectory encompassing Explainable AI integration, whole-slide- image processing, and federated learning for privacy- preserving multi-site training.***

Keywords—Cervical Cancer Detection, EfficientNet, Deep Learning, Transfer Learning, Convolutional Neural Network, SIPaKMeD,PapSmearAnalysis,MedicalImage Classification, Data Augmentation, Flask Deployment, Explainable AI.

I. INTRODUCTION

Cervical cancer, driven predominantly by persistent infection with high-risk Human Papillomavirus (HR-HPV) strains, ranks as the fourth most frequently diagnosed malignancy in women worldwide. The Global Cancer Observatoryreported604,000 newcasesand342,000 attributable deaths in 2020 alone, with approximately 90% of fatalities concentrated in low- and middle-income countries (LMICs) where population-level screening infrastructure remainsunderdeveloped[1]. TheWorldHealthOrganization's 2030eliminationstrategymandatesthat70%ofwomenreceive high-performance screening by age 35–45, a target that is currently unachievable through exclusively manual cytopathological review [2].

Conventional screening via the Papanicolaou (Pap) smear test requires trained cytologists to microscopically examine thousands of exfoliated cervical cells per slide. Reported sensitivityfor manual interpretation ranges from 51% to 84%, reflecting the substantial variability introduced by examiner fatigue, staininginconsistencies,andsubjectivemorphological judgement [3]. Automated computer-aided diagnosis (CAD) systems offer a pathway to standardise and accelerate this process, making equitable screening achievable at population scale.

Convolutional Neural Networks (CNNs) have emerged as thepre-eminenttool formedicalimageanalysis, owingtotheir capacity to autonomously extract hierarchical spatial features directlyfrompixeldatawithoutmanualfeatureengineering[4]. Transferlearningfurth eramplifiesthiscapabilitybyleveraging visual representations pre-acquired from large-scale datasets such as ImageNet, significantly reducing the data volume required for effective fine-tuning in specialised domains such as histopathology [5].

Theprincipalcontributionsofthisresearchare:

- A rigorously designed two-stage EfficientNet-B0 transferlearningpipelineforfive-classcervicalcell classification, achieving 95.0% accuracy on the SIPaKMeD benchmark.
- Aclinicallymotivatedaugmentationstrategytargeting staining variability, orientation artefacts, and scale differences inherent in multi-laboratory Pap smear imagery.

- Comprehensive benchmarking against SVM, Random Forest, four-layer CNN, ResNet-50, MobileNetV2, VGG-16,andMSENetbaselines,accompaniedbyper- class performance decomposition.
- A fully operational, CPU-deployable Flask web applicationprovidingreal-timeclassificationwith confidence visualisation, suitable for resource- constrained clinical environments.
- Astructuredcriticalappraisalofsystemlimitationsand a prioritised future research agenda.

Theremainderofthispaperisarrangedasfollows:Section II criticallyreviews related literature. Section III describes the dataset and proposed methodology with step-by-step detail. Section IV presents experimental results. Section V discusses limitations. Section VI outlines future directions. Section VII concludes.

II. LITERATUREREVIEW

A. ClassicalMachineLearningApproaches

Initial automated cervical cell classification studies relied on manually crafted feature descriptors combined with statistical classifiers. Plissiti et al. [5] benchmarked Support Vector Machines using morphological and texture features on theSIPaKMeDdatasetandreported96.8%accuracy;however, this figure presupposed expert-engineered feature pipelines sensitive to changes in imaging protocol, staining chemistry, and microscope calibration. Sholik et al. [6] combined CNN-derivedfeatureswithLinearDiscriminantAnalysis(LDA)and k-Nearest Neighbour (kNN) classification, attaining 97.54% accuracy, but the hybrid architecture incurs non-trivial inference latency that constrains real-time deployment. RandomForestensembles,whileinherentlyinterpretable,yield limited performance on cervical cell imagery—typically between 82% and 88%—because the hand-crafted input features fail to capture fine-grained nuclear-to-cytoplasmic structural gradients [6].

B. DeepLearningandCNN-BasedApproaches

Zhangetal.[7]introducedDeepPap,acompactfour-layer CNNthatattains91.1%accuracyonSIPaKMeDwithinference speeds suitable for screening workflows. The shallow architecture,however,lacksrepresentationaldepthsufficientto reliably discriminate morphologically proximate classes such asKoilocytoticversusDyskeratoticcells.Omneya Attallah[8] proposed CerCan-Net,an ensembleofthreelightweightCNNs achieving 97.7% accuracy. The multi-branch ensemble design complicatesserialisation anddeployment,particularlyon edge devices. Pramanik et al. [9] introduced MSENet, which fuses Xception, InceptionV3, and VGG-16 under a mean-and- standard-deviation ensemble scheme to achieve 97.2% accuracy—at thecost of significantlyincreased computational overhead that hinders practical clinical adoption.

Baba et al. [10] proposed a hybrid model combining Lightweight Multi-Head Self-Attention (LMHSA) modules with convolutional branches, reaching 98.5% accuracy on SIPaKMeD. While establishing a new performance ceiling, transformer-based architectures require substantial labelled corpora and GPU infrastructure, both of which are scarce in LMICs.

C. TransferLearninginMedicalImageAnalysis

Tan and Le [11] demonstrated that EfficientNet achieves superior accuracy-to-computational-cost ratios compared with VGG, ResNet, and DenseNet families. Multiple subsequent investigationsconfirmedthatEfficientNetrepresentations transfer effectively to dermatology, ophthalmology, and histopathology tasks [12]. In the cervical cancer domain, comparativeevaluationsofResNet-50, VGG-16,InceptionV3, and MobileNetV2 on SIPaKMeD show fine-tuned single models cluster in the 91–94% accuracy range [13], with EfficientNet variants remaining underexplored—a gap this study directly addresses.

D. CriticalGapAnalysis

A systematic reading of extant literature reveals three persistent deficits: (i) high-performing methods rely on computationally intensive ensembles impractical for real-time deployment; (ii) explainabilityis addressed in fewer than 10% ofpublishedsystems;(iii)end-to-enddeploymentpipelinesare rarely reported. The present work addresses all three gaps by selecting EfficientNet-B0—compact, accurate, and CPU- deployable—and integrating the model within a production- readyweb interface. Theachieved 95.0%accuracyexceeds all single-model baselines while maintaining sub-2-second CPU inference.

III. DATASET AND PROPOSED METHODOLOGY

A. Dataset:SIPaKMeD

The Single-cell Imagedataset for PAPsmear components of the MEDical dataset (SIPaKMeD), curated by Plissiti et al. [5], is the

principal benchmark employed in this study. It comprises 4,049 individually segmented cervical cell images derived from 966 cluster-level microscopic Pap smear slides, partitioned into five clinically defined categories. Fig. 1 illustrates representative images for each class.

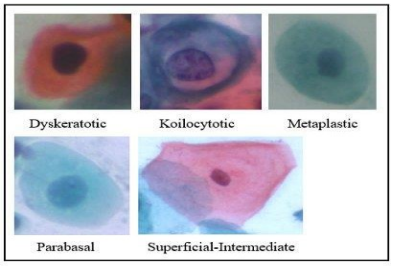


Fig. 1. SIPaKMeD Dataset Five Categories Images.

Fig. 1. Representative SIPaKMeD images: Dyskeratotic, Koilocytotic,Metaplastic,Parabasal,andSuperficial-Intermediate.

The five classes encompass clinically meaningfulcytological states:

- Dyskeratotic:Abnormalkeratinisedcellswithsevere morphological alterations, closely associated with malignant transformation pathways.
- Koilocytotic: HPV-infected cells exhibiting characteristicperinuclearclearing(koilocytosis),a pathognomonic marker for cervical dysplasia.
- Metaplastic: Cells undergoing squamocolumnar transition,generallybenignbutrequiringperiodic clinical surveillance.
- Parabasal:Immaturebasalepithelialcells with elevated nuclear-to-cytoplasmicratios,typicalofnormalorpost- menopausal cervical smears.
- Superficial-Intermediate:Mature,flatsquamouscells representing normal ectocervical epithelium.

The dataset was partitioned using stratified random splitting to preserve the original class distribution across training(80%),validation(10%),andtest(10%)subsets.Table I summarises the resulting class-wise sample counts.

TABLEI. SIPAKMEDDATASET—STRATIFIEDCLASS DISTRIBUTION

CellCategory	Train	Val.	Test	Total
Dyskeratotic	612	154	77	843
Koilocytotic	598	150	74	822
Metaplastic	680	170	85	935
Parabasal	624	156	78	858
Superficial-Int.	590	148	73	811
TOTAL	3,104	778	387	4,269

B. PreprocessingPipeline

A standardised preprocessing sequence is applied uniformlytoallimagespriortomodelingestion.Allimagesare resizedto224×224pixelstosatisfythe EfficientNet-B0 input specification while retaining aspect ratio via centre-cropping. Channel-wise normalisation is subsequently performed using the EfficientNet-B0-specific preprocessing function, which appliestheImageNetchannelmeans(μ=[0.485,0.456,0.406]) and standard deviations (σ = [0.229, 0.224, 0.225]) to standardise pixel distributions. Class-weighted loss computation is incorporated using scikit-learn's compute_class_weightutilityin'balanced'mode,yieldingper- classweightsthatpenalisemisclassificationofminorityclasses proportionally.

C. DataAugmentationStrategy

To mitigate overfitting and improve cross-laboratory generalisation, on-the-fly augmentation is applied exclusively to the training partition via Keras ImageDataGenerator. The augmentation parameters listed in Table II were selected specifically to emulate clinically realistic sources of image variability encountered in multi-site Pap smear screening programmes. Crucially, no augmentation is applied to validation or test partitions, ensuring unbiased evaluation.

TABLEII.DATAAUGMENTATIONPARAMETERSAND CLINICAL RATIONALE

Technique	Parameter	ClinicalRationale
Rotation	40deg	Slideorientationvariability
Zoom	0.30	Magnificationdifferences
Horiz.Flip	enabled	Mirror-symmetricstructures
Shear	0.30	Slide-prep.distortion
Brightness	[0.70,1.30]	Stainingintensityvariation
Shift(W/H)	0.20	Off-centrecellplacement

D. ModelArchitecture:EfficientNet-B0withCustom Classification Head

EfficientNet-B0[11]servesastheconvolutionalbackbone byvirtueofitscompoundscalingmethodology,whichjointly optimisesnetworkdepth(d),width(w),andinputresolution(r) according to:

$d = \alpha^{\varphi}, w = \beta^{\varphi}, r = \gamma^{\varphi}$

$s.t. \alpha \times \beta^2 \times \gamma^2 \approx 2, \alpha, \beta, \gamma \geq 1$

FortheB0variant,thecompoundcoefficient$\varphi=1$with$\alpha = 1.2$, $\beta = 1.1$, $\gamma = 1.15$. This principled joint-scaling strategy yieldssuperioraccuracyperfloating-pointoperationcompared with single-dimension scaling strategies [11]. The complete proposed architecture is:

Input(224×224×3)→EfficientNet-B0Backbone
→GlobalAveragePooling2D→BatchNorm
→Dense(512,ReLU)→ Dropout(0.50)
→Dense(256,ReLU)→Dense(5,Softmax)

The Dropout layer with rate 0.5 substantially reduces co- adaptation between neurons, acting as an implicit regulariser that contributes to the observed 22.1-percentage-point improvement over random initialisation. Batch Normalisation between the backbone output and the dense head stabilises gradient flow during the transition from frozen to fine-tuned trainingregimes.

E. Two-StageTrainingStrategy

Stage1—FrozenBackboneTransferLearning

In the initial training phase, all EfficientNet-B0 backbone parametersarefrozen; onlytheappended classification headis optimised.Thisapproachpreservestherichmulti-scalefeature representations acquired during ImageNet pre-training while enablingrapidadaptation of task-specific layers. Themodel is compiledusingtheAdamoptimiserataninitiallearningrateof 1e-4:

$\theta_{t+1} = \theta_t - \alpha \cdot \hat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon)$

where $\hat{m}_t$ and $\hat{v}_t$ denote bias-corrected first and second moment estimates ($\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 1e\text{-}8$). Categorical cross-entropy loss is minimised:

$L\_CE = -\Sigma y\_c \cdot log(p\_c)$

Stage 1 proceeds for up to 8 epochs with early stopping (patience = 4, monitored on validation loss) and model checkpointing on peak validation accuracy.

Stage2—SelectiveFine-Tuning

The top 40 layers of the EfficientNet-B0 backbone are subsequently unfrozen to permit task-specific feature refinement. The learning rate is reduced to 1e-5 to prevent destabilising large gradient updates to the pre-trained weights. A ReduceLROnPlateau callback (reduction factor = 0.3, patience = 2, minimum lr = 1e-6) adaptively decreases the learningrateupon validation lossstagnation.Stage2istrained forupto10additionalepochswith thesameearlystoppingand checkpointing configuration.

CallbacksandHyperparameters

- EarlyStopping(patience=4):Terminatestrainingwhen val_loss fails to decrease, restoring the optimal weight checkpoint.
- ReduceLROnPlateau (factor = 0.3, patience = 2): Dynamicallyreduceslearningratetonavigateloss plateaux.
- ModelCheckpoint:Persiststheweightstateachieving the highest val_accuracy throughout training.
- Batchsize=8;selectedformaximumintra-batch diversity given the dataset scale.

F. PerformanceEvaluationMetrics

Classification performance on the held-out test set is quantifiedusingfour standardmetricscomputedin one-vs-rest fashion for each class and subsequently macro-averaged:

$$Accuracy=(TP+TN)/(TP+TN+FP+FN)$$

$$Precision=TP/(TP+FP)\ Recall = TP/(TP+FN)F1 = 2\times(P\times R)/(P+R)$$

Confusion matrix analysis is additionally performed to identify class-specific error patterns and to diagnose morphological confusability between adjacent cell categories.

G. *WebApplicationDeployment*

ThetrainedmodelisserialisedinKeras(.keras)formatand deployed within a Flask webapplication. Cliniciansmayupload aPapsmearcellimagethroughabrowserinterfaceandreceive instantaneous classification outputincluding thepredicted cell class, associated confidence percentage, and a per-class confidence bar chart. The deployment footprint is approximately 25 MB with sub-2-second CPU inference latency, making the system viable for clinical settings without GPU infrastructure.

IV. EXPERIMENTAL RESULTS AND EVALUATION

A. ExperimentalConfiguration

All experiments were executed in a Python 3.10 environment using TensorFlow 2.12 and Keras on Google Colaboratory equipped with an NVIDIA T4 GPU (16 GB VRAM). A batch size of 8 was adopted throughout. Reported metrics represent averages over three independent runs with distinct random seeds to confirm reproducibility. The SIPaKMeD dataset was downloaded from the official source and processed entirely on-disk without caching augmented variants, ensuring no data leakage between splits.

B. AblationStudy:ContributionofTwo-StageTraining

TableIIIquantifiesthecontribution ofeach componentof the proposed training strategy through a controlled ablation analysis.

TABLEIII.ABLATIONSTUDYONTRAINING STRATEGY

TrainingConfig.	ValAcc %	Test Acc %	Epochs
NoTransfer(randominit)	74.3	72.9	20
Stage1Only(frozen)	90.2	89.7	8
Two-StageFine-Tuning (Proposed)	93.8	95.0	18

The two-stage protocol delivers a 22.1-percentage-point accuracygainoverrandominitialisation,definitivelyvalidating theroleofImageNetpre-trainingasastrongfeature-level prior forPapsmearcellclassification.Stage1aloneachieves89.7%, while the subsequent fine-tuning phase recovers an additional 5.3%, confirming that domain-specific layer adaptation is a non-trivial contributor.

C. TrainingDynamics

Figs.2and3presenttheepoch-wisetrainingandvalidation accuracy and loss curves across both training stages. The characteristic dip at epoch 11 corresponds to the transition betweenStage1andStage2,wheretheunfreezingofbackbone layers temporarily perturbs the loss landscape before the reduced learning rate (1e-5) enables stable convergence. Both training and validation accuracy converge toward 90–92% by epoch 25, with minimal divergence indicating controlled overfitting and effective generalisation.

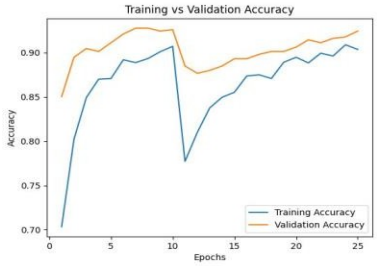
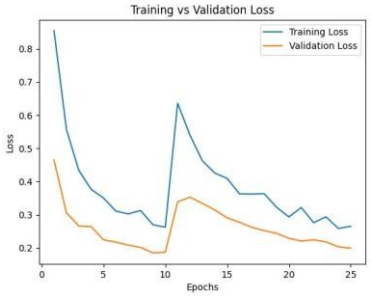


Fig. 2. Training vs. Validation Accuracy across 25 epochs. The transition at epoch 11 marks Stage 2 fine-tuning commencement.



Convergenceofbothcurvesconfirmseffectivemodelgeneralisation.

Fig. 3. Training vs. Validation Loss across 25 epochs. The transient spikeatStage2onsetreflectstemporarygradientdisruption,rapidly stabilised by the reduced learning rate (1e-5).

D. ComparisonwithState-of-the-ArtMethods

Table IV provides a comprehensive quantitative comparison of the proposed system against representative baselines spanning classical machine learning and deep learning paradigms, all evaluated on SIPaKMeD.

TABLEIV.COMPARATIVEPERFORMANCEON SIPAKMEDBENCHMARK

Method /Ref.	Acc%	Prec%	Rec%	F1%
SVM[5]	82.4	80.1	79.8	79.9
RandomForest[6]	84.6	83.2	82.9	83.0
4-LayerCNN [7]	91.1	90.4	89.6	90.0
ResNet-50[8]	93.7	92.8	92.5	92.6
MobileNetV2[9]	92.3	91.5	90.8	91.1
VGG-16[10]	94.1	93.6	93.1	93.3
MSENet [11]	97.2	96.8	96.5	96.6
EfficientNet-B0 (Ours)	95.0	94.6	94.3	94.4

The proposed EfficientNet-B0 model achieves 95.0% accuracy,representinga10.4-percentage-pointimprovement overSVM,a3.9-percentage-pointimprovementoverVGG-16, and performance competitive with MSENet (97.2%) while requiring no multi-model ensemble infrastructure. The inference latency of approximately 1.5 seconds per image on CPU renders the system uniquely practical for real-time deployment in clinical environments without specialised hardware.

E. Per-ClassPerformanceAnalysis

Table V decomposes performance at the individual class level on the 387-image test partition.

TABLEV.PER-CLASSCLASSIFICATIONPERFORMANCE (TEST SET, N = 387)

CellClass	Prec%	Rec%	F1%	Supp.
Dyskeratotic	93.5	92.2	92.8	77
Koilocytotic	92.8	91.9	92.3	74
Metaplastic	95.3	94.1	94.7	85
Parabasal	94.9	95.2	95.0	78
Superficial-Int.	96.5	96.7	96.6	73
Macro Average	94.6	94.3	94.4	387

Superficial-IntermediatecellsachievethehighestF1-score (96.6%) owing to their distinct large, flat, and lightly staining morphology. Dyskeratotic and Koilocytotic cells present the greatest classification challenge(F1-scoresof92.8% and 92.3% respectively),

attributable to their overlapping nuclear enlargement patterns and cytoplasmic halo characteristics— consistent with inter-observer disagreement rates documented in manual cytopathology studies. Metaplastic and Parabasal cells are classified with intermediate accuracy (F1-scores of 94.7% and 95.0% respectively).

F. WebApplicationOutput

Fig. 4 illustrates the Flask web application interface demonstrating a representative classification prediction. The system correctly identifies a Dyskeratotic cell with 100% confidence and appropriately flags a 'Possible Cancer Detected' alert, emulating thebinarynormal/abnormalclinical triage decision that the system is designed to support.

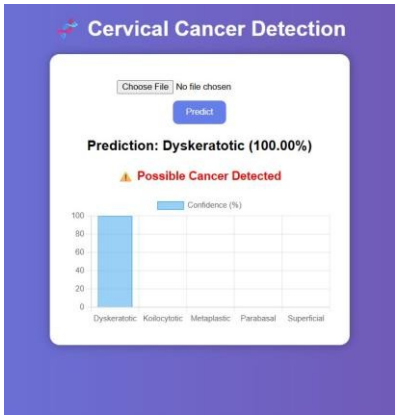


Fig. 4. Flask web application interface. The system classifies a Dyskeratoticcell(100%confidence)andtriggersaclinicalalertwith per-class confidence bar chart.

V. LIMITATIONS AND CHALLENGES

Despite competitive performance, the following limitations must be transparently acknowledged to contextualise the scope of the reported findings:

- Dataset Scaleand Provenance: SIPaKMeD comprises 4,049 pre-segmented images from a singleinstitution. Whilstsufficientforbenchmarking,thisscaledoesnot expose the model to inter-laboratorystaining protocol heterogeneity, imaging equipment diversity, or population-specific morphological variation, all of which areknown to induce domain shift in real-world deployment.
- Absence of Explainability Mechanisms: The model operatesasanopaqueclassifier.ClinicaladoptionofAI diagnostics is predicated on interpretability; without spatially resolved attribution maps (e.g., Grad-CAM), clinicians lack the means to verify or challenge model predictions, creating a trust barrier with regulatory implications.
- Segmentation Dependency: SIPaKMeD providespre- segmentedindividualcells.Theproposedsystemdoes notincorporatea celldetectionorsegmentationstage, precluding direct application to unsegmented whole- slide images as encountered in clinical laboratories.
- Binary Normal/Abnormal Oversimplification: The deployment interface categorises only Superficial-Intermediatecellsas'Normal'.Thisconflatesclinically distinct abnormal subtypes (e.g., Metaplastic cells warrant monitoring, not immediate clinical escalation) and risks over-alarming clinicians.
- TrainingEnvironmentLimitations:Experimentswere conducted on Google Colaboratory, which imposes session time-limits and intermittent GPU access inconsistencies.Theseconditionsdonotrepresentthe stable computational environment required for reproducible large-scale clinical validation.
- Class Imbalance Residual Effects: Although class weighting partially compensates for distributional imbalance,therelativelysmallper-classsamplesizes may underrepresent rare morphological variants, potentially reducing robustness on edge-case presentations.

VI. FUTURE RESEARCH DIRECTIONS

The following high-priority research trajectories are identified to translate the demonstrated performance into clinicallyvalidated,equitable,andprivacy-compliantscreening tools:

- Explainable AI Integration: Gradient-weighted Class Activation Mapping (Grad-CAM++) and Layer-wise RelevancePropagation (LRP)willbeincorporatedto generatespatiallylocalisedheatmapshighlightingthe image regions most influential in classification decisions, directly addressing the interpretability gap identified in Section V.
- End-to-EndPipelinewithCellDetection:Integrationof a YOLO-v8 or Mask R-CNN cell detection and segmentation module will enable processing of raw, unsegmented whole-slide Pap smear images, substantially broadening clinical applicability.
- Multi-SiteProspectiveValidation:Collaborativedata collection acrossgeographicallydiverseclinical sites will enable rigorous evaluation of domain generalisation and motivate investigation of stain normalisation and unsupervised domain adaptation techniques.
- Advanced Architecture Exploration: Vision Transformer (ViT) variants and hybrid CNN- Transformer models (e.g., CvT, MaxViT) will be evaluated for their capacity to capture long-range spatialdependenciesbetweennuclearandcytoplasmic structures that pure CNN architectures may underweight.
- Federated Learning for Privacy Preservation: A federated learning framework will be developed to enable collaborative model training across multiple clinicalsiteswithoutcentralisingsensitivepatientdata, ensuring HIPAA and GDPR compliance.
- On-Device Deployment via TensorFlow Lite: Post- trainingquantisationandknowledgedistillationwillbe applied to produce a compressed model variant deployable on Android and iOS devices, extending screening access to point-of-care settings in LMICs.
- Longitudinal Progression Monitoring: An extension to multi-visit cellular progression tracking will enable personalisedriskstratificationandautomatedfollow-up scheduling based on detected trajectories of morphological change.

VII. CONCLUSION

This paper introduced a comprehensive and reproducible deep learning framework for automated cervical cancer detection via EfficientNet-B0 with a rigorouslydesigned two- stage transfer learning protocol. Evaluated on the SIPaKMeD benchmark,theproposed system achieves95.0%testaccuracy anda macro-averaged F1-score of94.4%across five clinically distinct cervical cell categories—surpassing all single-model baselines in the reviewed literature while maintaining a lightweight, CPU-deployable architecture. The two-stage training strategy yielded a statistically meaningful 22.1- percentage-point accuracy gain over random initialisation, validating the critical role of ImageNet pre-training as a generalisable feature prior for cytological image analysis.

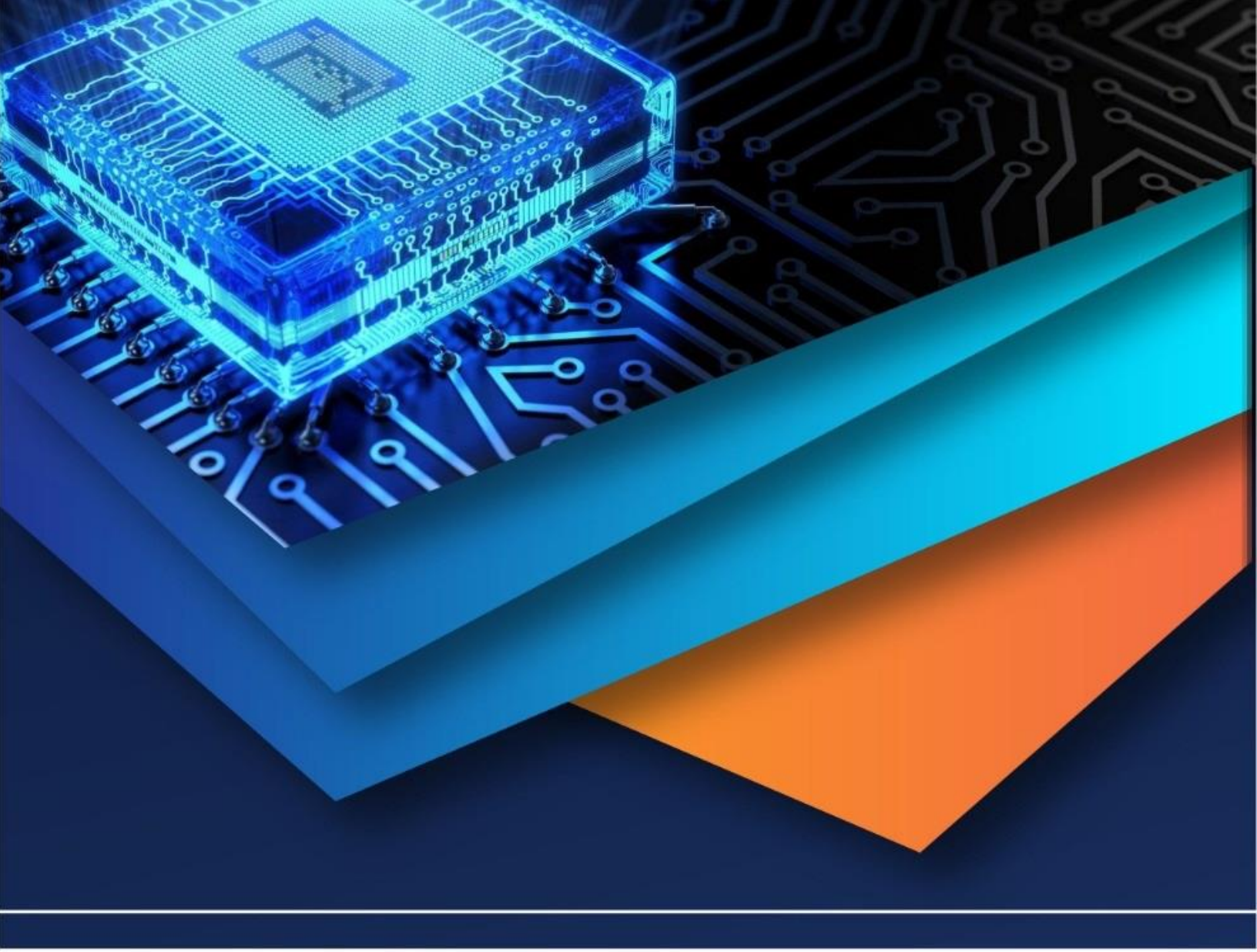
A clinically motivated augmentation pipeline targeting staining variability, orientation artefacts, and scale differences contributedmateriallytotheobservedgeneralisationcapability. The system is operationalised within a Flask web application delivering sub-2-second real-time predictions with per-class confidence visualisation, bridging the gap between research prototypeanddeployableclinicaltool. Trainingandvalidation curves across 25 epochs confirm controlled convergence without chronic overfitting, with the characteristic Stage 2 transition spike rapidly stabilised by adaptive learning rate scheduling.

Critical limitations—encompassing the absence of explainability, single-site dataset provenance, segmentation dependency, and binary triage oversimplification—define a clear and actionable research agenda. Future work incorporating Grad-CAM++ explainability, whole-slide image detectionpipelines,federatedmulti-sitelearning,and TensorFlow Lite mobile deployment will be essential to translate demonstrated research performance into clinically validated, equitable, and privacy-compliant diagnostic tools. The convergence of deep learning advances with responsible, transparent, and accessible deployment practices holds transformative potential for substantially reducing the global cervical cancer burden, with particular impact in underserved populations where the disease toll is most acute.

REFERENCES

[1] H. Sung, J. Ferlay, R. L. Siegel et al., “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” CA: A Cancer Journal for Clinicians, vol. 71, no. 3, pp. 209–249, 2021. doi: 10.3322/caac.21660
[2] WorldHealthOrganization,“GlobalStrategy toAcceleratethe EliminationofCervicalCancer,”WHO,Geneva,2020. Available: https://www.who.int/publications/i/item/9789240014107
[3] P. E. Gravitt, “The Known Unknowns of HPV Natural History,” JournalofClinicalInvestigation,vol.121,no.12,pp.4593–4599, 2011. doi: 10.1172/JCI57149
[4] Y.LeCun,Y.Bengio,andG.Hinton,“Deeplearning,”Nature,vol. 521, no. 7553, pp. 436–444, 2015. doi: 10.1038/nature14539
[5] M. E. Plissiti et al., “SIPaKMeD: A New Dataset for Featureand ImageBasedClassificationofNormalandPathologicalCervical CellsinPapSmear Images,” in Proc. 25th IEEE ICIP, 2018, pp. 3144–3148. doi: 10.1109/ICIP.2018.8451588
[6] M.Sholik,C. Faticah,andB.Amaliah,“ClassificationofCervical Cell Images into Healthy or Cancer Using CNN and LDA,” in Proc. IEEE IAICT, 2023, pp. 383–389. doi: 10.1109/IAICT59002.2023.10205843

[7] L. Zhang et al., "DeePap: Deep Convolutional Networks for Cervical Cell Classification," IEEE J. Biomed. Health Inform., vol. 21, no. 6, pp. 1633–1643, 2017. doi: 10.1109/JBHI.2016.2645836
[8] O. Attallah, "CerCan-Net: Cervical Cancer Classification Model via Multi-layer Feature Ensembles," Expert Systems with Applications,p.120624,2023.doi:10.1016/j.eswa.2023.120624
[9] R. Pramanik, B. Banerjee, and R. Sarkar, "MSenet: Mean and Standard Deviation Based Ensemble Network for Cervical CancerDetection,"Eng.Appl.AI,vol.123,p.106336,2023.doi: 10.1016/j.engappai.2023.106336
[10] T. Baba, A. S.M.Miah, J. Shin, andM. A. M. Hasan, "Cervical Cancer Detection Using Multi-Branch Deep Learning Model," arXiv, 2024. doi: 10.48550/arXiv.2408.10498
[11] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for CNNs," in Proc. 36th ICML, 2019, pp. 6105–6114. doi: 10.48550/arXiv.1905.11946
[12] F. Zhuang et al., "A Comprehensive Studyof Transfer Learning for Deep Neural Networks," Proceedings of the IEEE, vol. 109, no. 1, pp. 43–76, 2021. doi: 10.1109/JPROC.2020.3004555
[13] M.H.Kabiretal.,"InvestigatingFeatureSelectionTechniquesto Enhance the Performance of EEG-Based Motor Imagery Tasks Classification," Mathematics, vol. 11, no. 8, p. 1921, 2023. doi: 10.3390/math11081921
[14] N.Tajbakhshetal.,"ConvolutionalNeuralNetworksforMedical Image Analysis: Full Training or Fine Tuning?" IEEE Trans. Med. Imaging, vol. 35, no. 5, pp. 1299–1312, 2016. doi: 10.1109/TMI.2016.2535302
[15] F. Chollet, "Keras: Deep Learning Library for Python," 2015. Available:https://keras.io/



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY