ICARC 2024

Sabragamuwa University of Sri Lanaka

Interpretable Cervical Cell Classification: A Comparative Analysis

Niruthikka Sritharan[1], Nishaanthini Gnanavel[1], Prathushan Inparaj[1], Dulani Meedeniya[1], Pratheepan Yogarajah[2]

Department of Computer Science and Engineering, University of Moratuwa[1],
School of Computing, Engineering and Intelligent Systems, Ulster University[2]

Presenting Author: Niruthikka Sritharan

Introduction

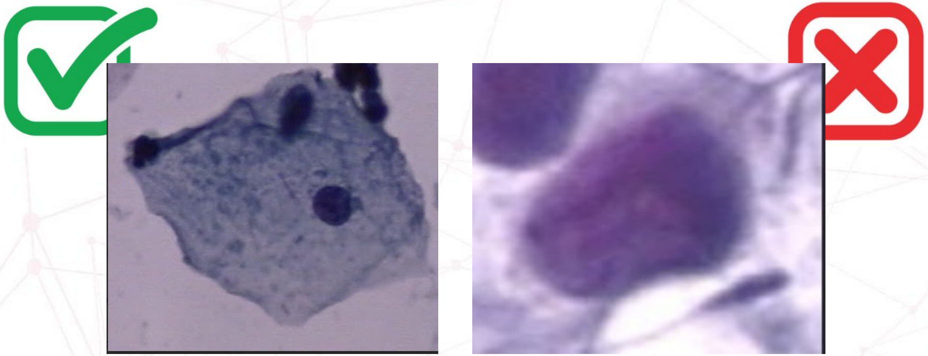
Cervical cancer
ranks as the
4th
most prevalent cancer among women [1].

In 2020, an estimated
341,831
women worldwide
died
from cervical cancer [1].

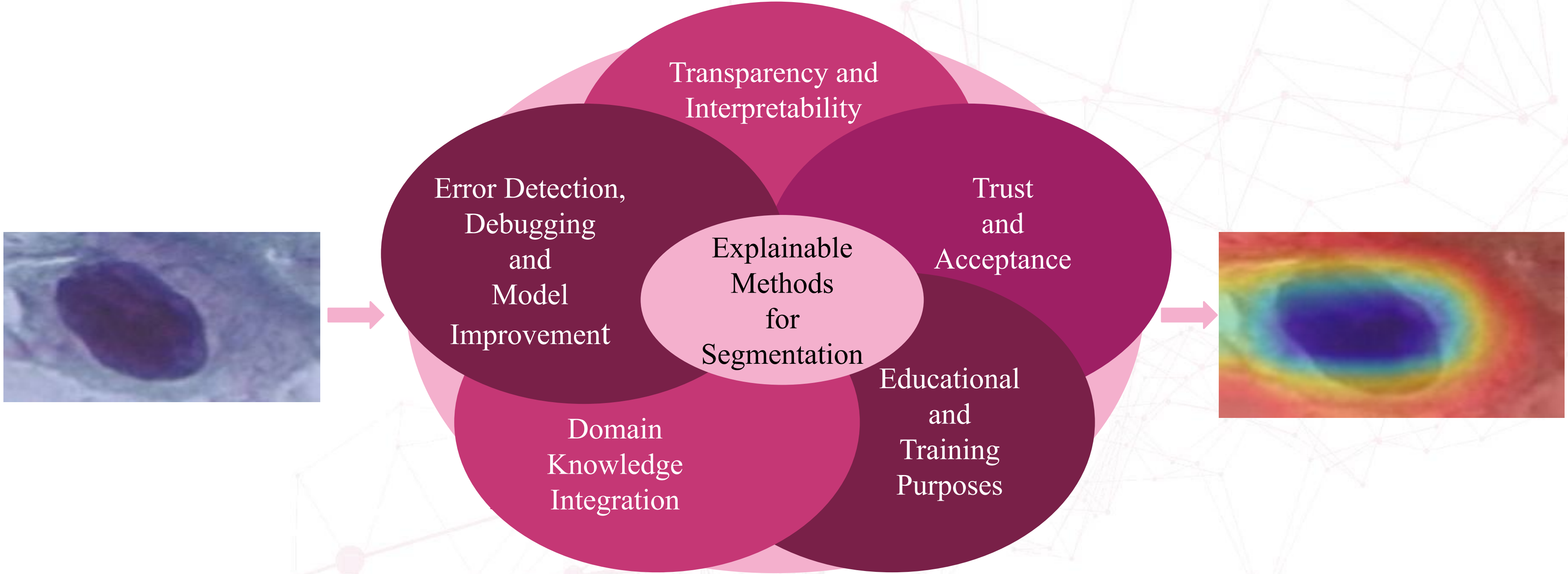
Incidence rates of cervical cancer dropped by more than
50%
from the mid-1970s
to the mid-2000s
due in part to an increase in
screening [1.]

Background

- Conventional Screening Methods:
  - ❑ Pap Smear Test
  - ❑ Liquid Based Cytology
- Limitations:
  - ❑ Time-consuming and laborious.
  - ❑ Prone to subjectivity causing unclear target boundaries [2].
  - ❑ A chance of 1 case to be missed in every 10 to 15 positive cases [3].



Motivation


Transparency and Interpretability
Error Detection, Debugging and Model Improvement
Trust and Acceptance
Explainable Methods for Segmentation
Domain Knowledge Integration
Educational and Training Purposes

Related Studies - Classification

Study	Techniques	Dataset	Performance
Diniz et al., 2021 [4]	• EfficientNet, MobileNet, XceptionNet • Ensemble Architecture	CRIC Searchable Image Database	A = 96%, P = 96%, R = 96%, S = 96%, F = 96%
Rahaman et al., 2021 [5]	• VGGNet, XceptionNet, ResNet • Hybrid Deep Feature Fusion network	SIPaKMeD	A = 99.85%, P = 100%, R = 100%, F = 99.8%
		Herlev	A = 98.91%, P = 99.5%, R = 98.0%, F = 98.5%
Alquran et al., 2022 [6]	• Cervical Net - A novel DL structure, Shuffle Net • Principal Component Analysis (PCA) and Canonical Correlation Analysis (CCA)	SIPaKMeD	A = 99.1%
Chowdary et al., 2023 [7]	• VGGNet, CaffeNet • Bag-of-features, linear-binary-patterns • PCA	Herlev	A = 98.39%, R = 98.97%, S = 97.65%
		SIPaKMeD	A = 99.16%, 99.15%, S = 99.75%

A = Accuracy, P = Precision, R = Recall, S = Specificity, F = F1 Score

Related Studies - Explainable Artificial Intelligence (XAI)

Study	Techniques	Dataset	Metrics Used
Pitroda et al., 2021 [8]	• Layer-wise Relevance Propagation (LRP) • Deep Taylor Decomposition (DTD) • Guided Backpropagation (GB) • Local Interpretable Model Agnostic Explanation (LIME)	Chest X-Ray images	Image entropy, Pixel flipping metric
Kaur et al., 2022 [9]	• Gradient-weighted Class Activation Mapping (GradCAM)	CT scan images	
Bhandari et al., 2023 [10]	• Shapley Additive Explanation (SHAP) • LIME	Kidney CT images	

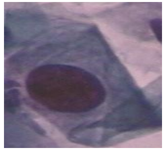
Novelty

- Utilizing XAI techniques in cervical cell classification.
- Providing a comprehensive qualitative analysis of the applied XAI techniques.
- Employing quantitative metrics on XAI techniques applied to the cervical classification task.

Methodology

Process Overview

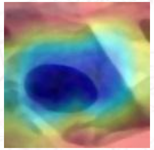
Input Data
Herlev Dataset [11]

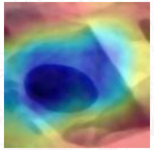


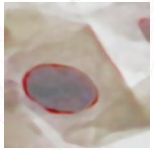
Data Preprocessing
- **Image Preprocessing**
  - Median filter
  - Histogram equalization and normalization
- **Image Augmentation**
  - Affine Transformations
  - Noise Additions
  - Color Shuffling

Supervised Binary Classification
- VGG16 [12]
- XceptionNet [13]
- EfficientNet [14]

XAI Methods
- GradCAM [15]
- GradCAM++ [16]
- LRP [17]



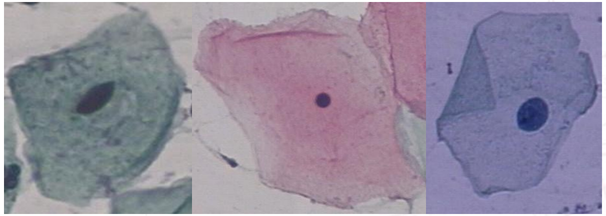




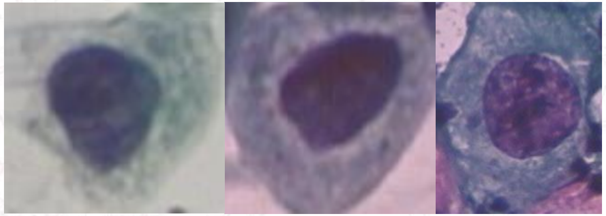
Evaluation
- **Qualitative Evaluation**
- **Quantitative Evaluation [8]**
  - Image Entropy
  - Pixel Flipping Performance Metric [18]

Dataset

- The dataset used for the study is Herlev dataset.
- It originally consists of 7 classes.
- For the scope of the study, we reclassified the dataset into 2 classes.
  - Normal class - Images without cancer cells
  - Abnormal class - Images with cancer cells



Normal Class



Abnormal Class

Image Preprocessing

- All images are resized to same size which is 224x224 pixels.
- Used median filter for noise reduction in images.
- Used histogram equalization and normalization to enhance the contrast [6].

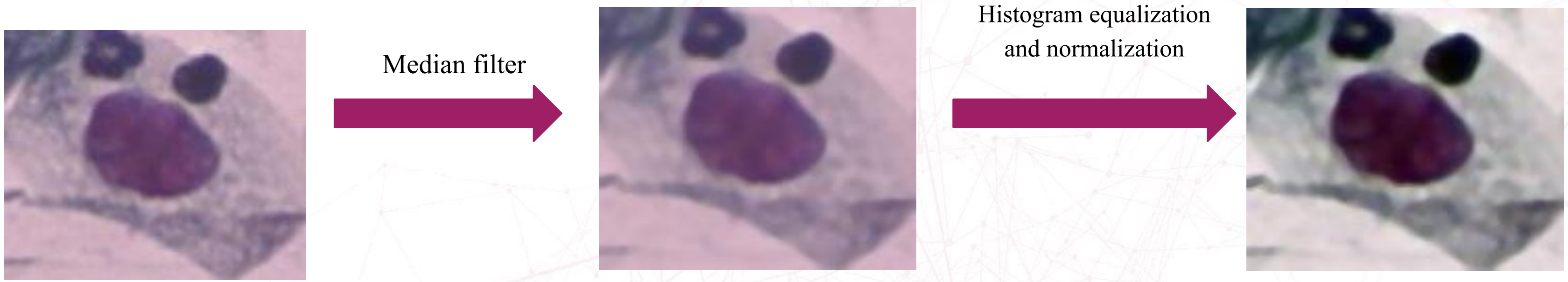
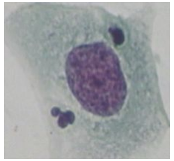
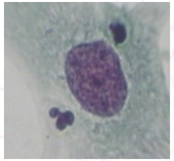
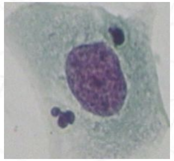
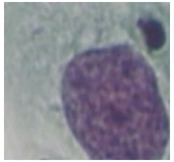
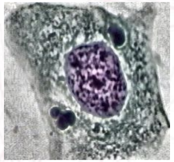
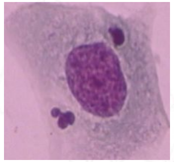
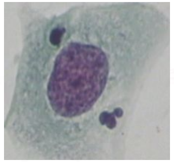
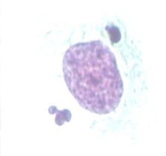


Image Augmentation

Original		**Optical distortion**, Grid distortion, Elastic transformations		Addition of **gaussian noise**, gaussian blur	
Translation, Scaling, **Rescaling of a random part**		**CLAHE**, image histogram equalization		Conversion to grayscale, shuffling color channel, addition of color jitter, **shifting RGB intensities**	
Vertical flip, **Horizontal flip**, Rotation by a random angle from 0 to 180 degrees		**Random modification of brightness, contrast of the image**		**Sharpening the image**	

Classification Models

Baseline	Hyperparameters		
	Batchsize	Learning Rate	Epochs
VGGNet	8	0.001	10
XceptionNet	16	0.001	10
EfficientNet	16	0.001	10

XAI Techniques

- GradCAM

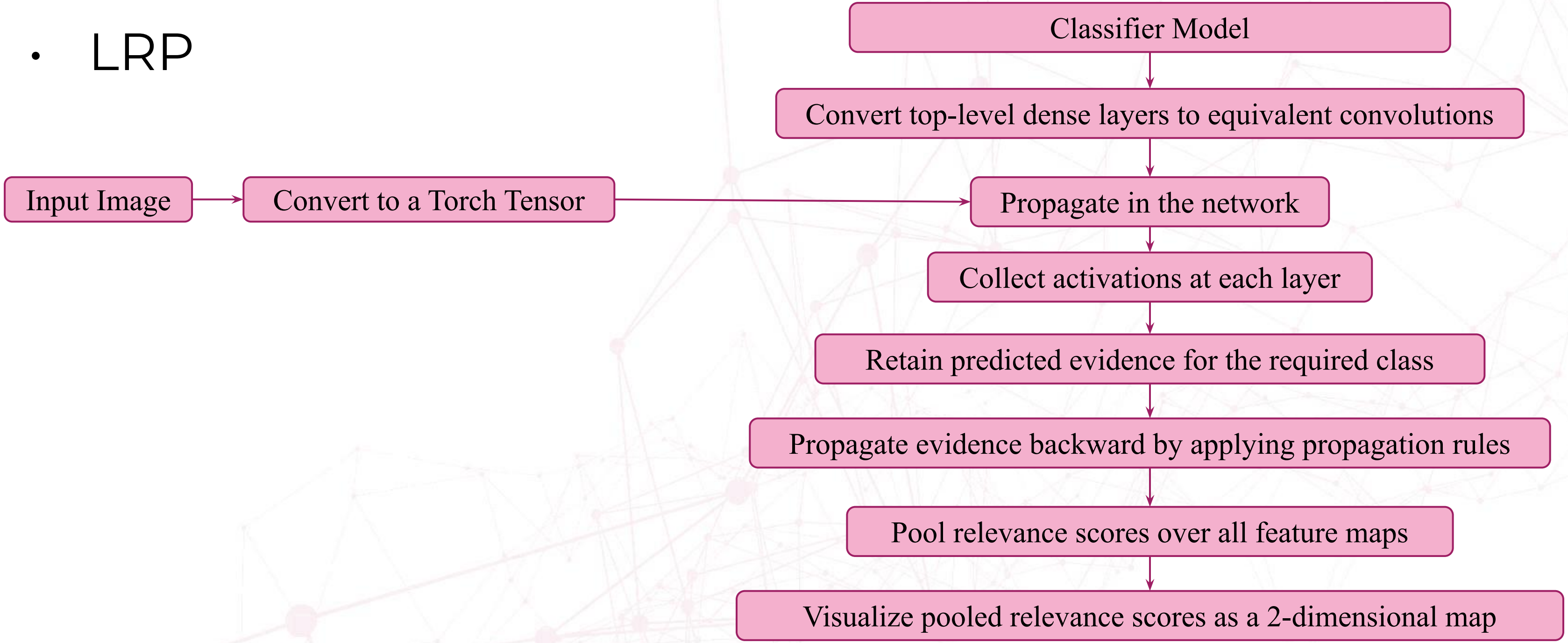
Input Image → Forward Pass → Backpropagation → Weighted Activation Maps → Heatmap Generation

- GradCAM++

Input Image → Forward Pass → Backpropagation → First Order Gradients Calculation → Second Order Gradients Calculation → Weighted Activation Maps → Heatmap Generation

XAI Techniques

- LRP



XAI Techniques

- LRP Ruleset

Original VGG Layer	Conv 3x3 (64) Conv 3x3 (64) MaxPool	Conv 3x3 (128) Conv 3x3 (128) MaxPool	Conv 3x3 (256) Conv 3x3 (256) Conv 3x3 (256) MaxPool	Conv 3x3 (512) Conv 3x3 (512) Conv 3x3 (512) MaxPool	Conv 3x3 (512) Conv 3x3 (512) Conv 3x3 (512) MaxPool	Linear (25088, 4096) Linear (4096, 4096) Linear (4096, 2)
Modified Layer	- - AvgPool (2)	- - AvgPool (2)	- - - AvgPool (2)	- - - AvgPool (2)	- - - AvgPool (2)	Conv 7x7 (512) Conv 1x1 (4096) Conv 1x1(2)
LRP Rule	z^B rule LRP ɣ LRP ɣ	LRP ɣ LRP ɣ LRP ɣ	LRP ɣ LRP ɣ LRP ɣ LRP ɣ	LRP ε LRP ε LRP ε LRP ε	LRP ε LRP ε LRP ε LRP ε	LRP 0 LRP 0 LRP 0

Evaluation Metric

Classification Models →

- Accuracy
- Precision
- Recall
- F1 score

XAI methods →

- Image entropy
- Pixel flipping performance metric

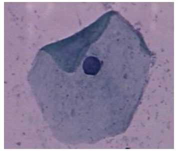
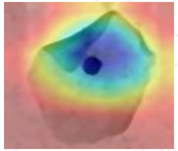
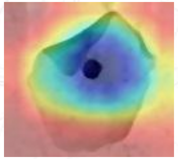
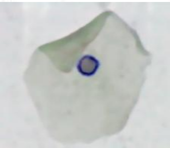
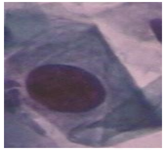
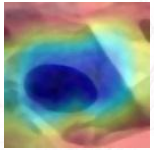
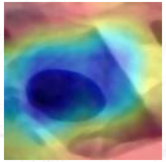
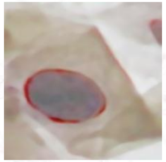
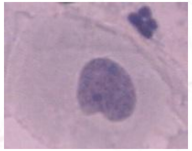
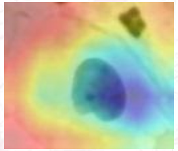
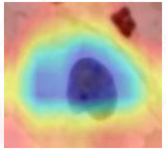
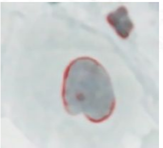
Classification Performance

Baseline	Accuracy	Precision	Recall	F1 Score
VGGNet	0.92	0.93	0.92	0.91
XceptionNet	0.86	0.88	0.86	0.84
EfficientNet	0.88	0.90	0.88	0.87

Performance of XAI Techniques

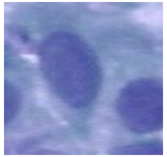
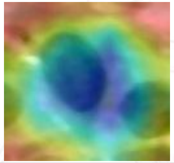
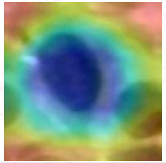
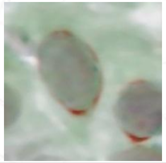
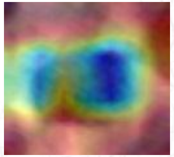
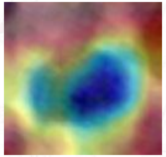
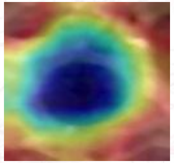
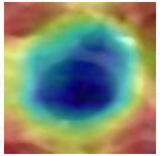
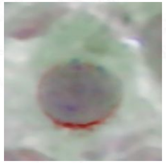
Qualitative Analysis

- Explanations for Best Predictions

No.	Original Image	GradCAM	GradCAM++	LRP
Image-1				
Image-2				
Image-3				

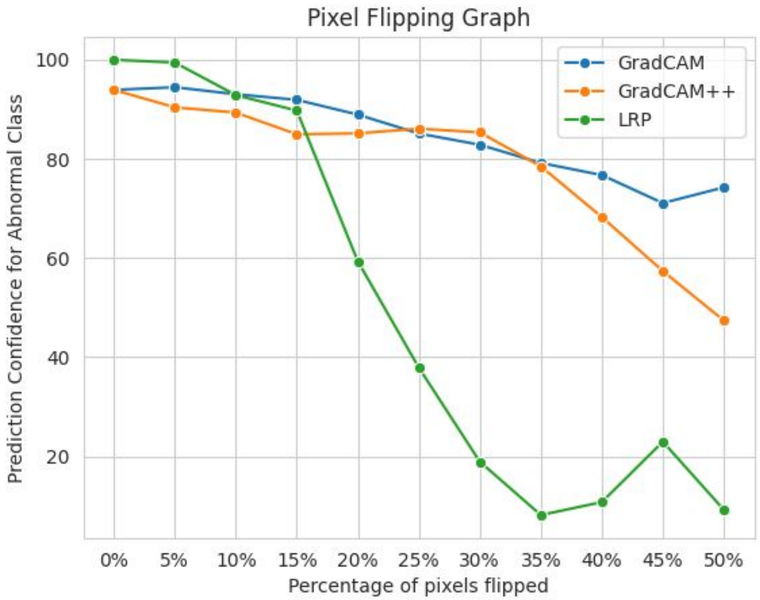
Qualitative Analysis

- Explanations for Worst Predictions

No.	Original Image	GradCAM	GradCAM++	LRP
Image-4				
Image-5				
Image-6				

Quantitative Analysis

XAI Technique	Mean Image Entropy
GradCAM	2.6567
GradCAM++	5.3688
LRP	2.4849



Conclusion

- Cervical cell image classification achieved 91.94% accuracy using VGG16-based CNN models.
- XAI techniques (GradCAM, GradCAM++, LRP) explained model decisions, emphasizing nuclei and cytoplasm as key indicators.
- LRP had lower complexity, highlighting the most relevant region for the classifier's decision.
- Quantitative comparisons favored LRP for its efficiency in identifying crucial features with lower entropy and steeper confidence drop.

References

[1] Cancer.net, "Cervical Cancer - Statistics," Cancer.net, Jun. 26, 2012. https://www.cancer.net/cancer-types/cervical-cancer/statistics

[2] "Cervical cancer: Symptoms, causes, stages, and treatment," *www.medicalnewstoday.com*, Sep. 27, 2021. https://www.medicalnewstoday.com/articles/what-you-need-to-know-about-cervical-cancer

[3] Y. Fan, Z. Tao, J. Lin, and H. Chen, "An encoder-decoder network for automatic clinical target volume target segmentation of cervical cancer in ct images," *International Journal of Crowd Science*, vol. 6, no. 3, pp.111–116, 2022

[4] D. N. Diniz, M. T. Rezende, A. GC Bianchi, C. M. Carneiro, E. JS Luz, G. JP Moreira, D. M. Ushizima, F. NS de Medeiros, and M. JF Souza, "A deep learning ensemble method to assist cytopathologists in pap test image classification," *Journal of Imaging*, vol. 7, no. 7, p. 111, 2021.

[5] M. M. Rahaman, C. Li, Y. Yao, F. Kulwa, X. Wu, X. Li, and Q. Wang, "Deepcervix: A deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques," *Computers in Biology and Medicine*, vol. 136, p. 104649, 2021.

References

[6] H. Alquran, M. Alsalatie, W. A. Mustafa, R. A. Abdi, and A. R. Ismail, "Cervical net: A novel cervical cancer classification using feature fusion," *Bioengineering*, vol. 9, no. 10, p. 578, 2022.

[7] G. J. Chowdary and P. Yogarajah, "Nucleus segmentation and classification using residual se-unet and feature concatenation approach incervical cytopathology cell images," *Technology in Cancer Research & Treatment*, vol. 22, p. 15330338221134833, 2023.

[8] V. Pitroda, M. M. Fouda, and Z. M. Fadlullah, "An explainable ai model for interpretable lung disease classification," in *IEEE International Conference on Internet of Things and Intelligence Systems (IoTaIS)*. Bandung, Indonesia: IEEE, 2021, pp. 98–103.

[9] A. Kaur, G. Dong, and A. Basu, "Gradxcepunet: Explainable ai based medical image segmentation," in *International Conference on Smart Multimedia*. Marseille, France: Springer, 2022, pp. 174–188.

[10] M. Bhandari, P. Yogarajah, M. S. Kavitha, and J. Condell, "Exploring the capabilities of a lightweight cnn model in accurately identifying renal abnormalities: Cysts, stones, and tumors, using lime and shap," *Applied Sciences*, vol. 13, no. 5, p. 3125, 2023.

References

[11] J. Jantzen, J. Norup, G. Dounias, and B. Bjerregaard, "Pap-smear benchmark data for pattern classification," *Nature inspired smart information systems (NiSIS 2005)*, pp. 1–9, 2005.

[12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[13] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[14] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in I*nternational conference on machine learning*. PMLR, 2019, pp. 6105–6114.

[15] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.

[16] A. Chattopadhay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks," in *IEEE winter conference on applications of computer vision (WACV)*. IEEE, 2018, pp. 839–847.

References

[17] S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. M¨uller, and W. Samek, "On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation," *PloS one*, vol. 10, no. 7, p.e0130140, 2015.

[18] J. Kauffmann, K.-R. M¨uller, and G. Montavon, "Towards explaining anomalies: a deep taylor decomposition of one-class models," *Pattern Recognition*, vol. 101, p. 107198, 2020.

Thank You