University of Ibadan
Journal of Science and Logics in ICT Research

Development of an AI Model for Automated Cervical Cancer Screening Based on Cytology Whole Slide Images

[1] Adeoye R. A., and [2] Adeyemo, A. B.

University of Ibadan, Ibadan
[1] *adeoye.richard.ar@gmail.com*, [2] *sesanadeyemo@gmail.com*

Abstract

Cervical cancer (CC) is the fourth leading cause of cancer-related deaths among women globally, posing a significant health challenge, especially in developing countries with limited healthcare resources. Traditional screening methods like the Pap smear, despite their success in early detection and reduction of CC incidence, face limitations such as labor-intensive analysis and subjective interpretation. Recent technological advancements in artificial intelligence (AI) and computer vision offer promising improvements in CC screening. This study focuses on developing and evaluating a convolutional neural network (CNN) for automated analysis of cervical cytology images from the SIPaKMeD database, which contains 4049 images categorized into five cell types: dyskeratotic, koilocytotic, metaplastic, parabasal, and superficial-intermediate. The proposed CNN architecture includes convolutional layers, MaxPooling, and Dropout layers to prevent overfitting, optimized using the Adam algorithm. Preprocessing steps such as color normalization, noise reduction, and artifact removal were applied to digitized slide images. The model's performance was assessed using accuracy and recall metrics. Results showed the CNN achieved high accuracy of 92% and sensitivity of 22%, with validation accuracy closely tracking training accuracy, indicating effective learning and generalization. The study demonstrates that Artificial Intelligence (AI) can significantly enhance CC screening accuracy and efficiency, addressing current limitations in traditional methods. Further optimization, including addressing class imbalance and exploring data augmentation techniques, is recommended to enhance the model's predictive capabilities and robustness in diverse clinical settings.

Keywords: *Cervical cancer, Artificial intelligence, Convolutional neural network, whole slide images, Automated diagnosis,*

1. Introduction

According to Kanavati [7], cervical cancer (CC) is the fourth most common cause of cancer-related deaths among women globally, and it continues to be a major global health concern. CC remains a problem even with improvements in screening methods and treatment choices, especially in developing nations with limited access to medical resources. Cao *et. al* [3] asserted that cervical cytology, often known as Pap smear, is the major screening approach that has been most successful in identifying precancerous changes in cervical cells. This has allowed for early intervention and a decrease in the incidence and mortality rates of cervical cancer. However, there are several inherent drawbacks to traditional screening approaches, such as the need for specialized individuals for interpretation, the need for labor-intensive manual analysis, and subjectivity in diagnosis, particularly in areas with inadequate healthcare infrastructure [5]. The difficulties in accurately performing CC screening are further exacerbated by differences in slide preparation methods and quality throughout laboratories.

Recent technological developments present intriguing ways to improve cervical cancer screening, especially in the areas of computer

Adeoye R. A., and 2 Adeyemo, A. B. (2024). Development of an AI Model for Automated Cervical Cancer Screening Based on Cytology Whole Slide Images. *University of Ibadan Journal of Science and Logics in ICT Research (UIJSLICTR),* Vol. 12 No. 1, pp. 13 – 21.

vision and AI [13]. There is a chance to increase the effectiveness and precision of CC identification by the automated analysis of cytology whole slide images (WSIs) using deep learning algorithms. Artificial Intelligence can be utilized to create systems that can more accurately detect aberrant and precancerous cells, which would enable prompt diagnosis and treatment.

Convolutional neural network (CNN) architecture development and evaluation for automated cervical cancer (CC) screening is the goal of this paper. In order to do this, a CNN architecture that is specifically designed to meet the needs of CC screening was created, implemented using the Python programming language, and its accuracy and recall were thoroughly assessed. In order to assure the construction of a scalable and efficient solution, the study also attempts to address important issues in automated screening systems, such as feature extraction, algorithm robustness, and slide scanning and preprocessing. This study intends to improve cervical cancer screening globally, especially in resource-constrained situations, by creating an accurate and effective AI model for Pap smear analysis using entire slide images with minimal pixel annotation. The use of AI into healthcare systems has promise for improving early detection, decreasing missed diagnoses, and eventually saving lives.

2. Related Works

A major global public health concern, cervical cancer (CC) is typified by abnormal cervix tissue growth, usually due to Human Papillomavirus (HPV) infection (United Nations, 2016). In women, colorectal cancer (CC) continues to be the primary cause of cancer-related fatalities despite improvements in screening methods, especially in areas with poor access to healthcare resources [7]. There are drawbacks to traditional screening techniques like Pap smears, such as interpretation subjectivity and dependence on trained staff [5]. The efficiency and accuracy of CC screening could be improved with the use of recent advances in AI. Early detection and prevention of cervical cancer are crucial, and screening for the disease usually include procedures like Pap smears and HPV tests [1]. AI-based solutions for CC screening have benefits including fewer time needs, less reliance on experts, and the removal of subjective biases [6]. When paired with HPV vaccination campaigns, AI integration has the potential to increase screening accuracy [12].

Images related to cervical cytopathology have been analyzed using a variety of deep learning (DL) techniques. These include CNN-based techniques like DenseNet, ResNet, and VGG16, which have demonstrated encouraging outcomes in differentiating between aberrant and normal cervical cells [5]. The potential of deep learning (DL) to transform CC screening has been demonstrated by the use of transfer learning approaches to improve model performance [9].

For automated CC screening, convolutional neural networks (CNNs)-based AI models have been created. Some of these models have outperformed conventional screening techniques in identifying cervical abnormalities, demonstrating high accuracy and specificity [2, 8]. It has been suggested that hybrid deep learning frameworks, like HDFCN, enhance classification accuracy by combining features from several CNN architectures [4]. But difficulties still exist, especially in terms of improving sensitivity to guarantee precise detection of positive cancer patients [11]. To sum up, AI techniques have enormous potential to improve cervical cancer screening's effectiveness and precision. To solve current issues and enhance screening results, more research and development in this area are needed.

3. Methodology

The main development platform for this paper was the Kaggle notebook and the Python programming language. Many Python libraries were used, including Keras, TensorFlow, NumPy, Matplotlib, and Python Image Library. An Intel Core i5-7200U CPU and 8GB RAM were among the technical specifications of the laptop used for development, which ran Windows 10 x64. Convolutional, pooling, and fully connected layers made up the sequential architecture used by the traditional CNN model. It began by processing full slide images into 224 by 224 pixel RGB image patches. 2D convolutional layers were used to understand low-level visual elements like edges and textures. After that, feature maps were progressively sampled down using max-pooling layers. Over several convolutional blocks, the number of filters was increased from 16 to 128 to allow the network to learn ever more complicated representations. The use of dropout regularization helped to lessen overfitting. In order to incorporate spatial information, flattened feature maps were processed through fully-connected layers with Rectified Linear Unit (ReLU) activations. The final layer, a softmax output layer, generated class probabilities that distinguished between abnormal/cancerous and normal/benign classes.

Table 1: SipakMed Dataset

Dataset	**SIPaKMeD Database**
Total Images	5,015
Single Cell Images	4,049
Cell Image Categories	5 (dyskeratotic, kilocytotic, metaplastic, parabasal, and superficial-intermediate)
Cluster Cell Images	966 (manually cropped)
Data Augmentation	The dataset was supplemented and labeled to increase diversity
Preprocessing	Images were preprocessed

Table 1 captures the key details about the SIPaKMeD database, including the total number of images, the breakdown of single cell and cluster cell images, the five categories of single cell images, the data augmentation and preprocessing steps applied to the dataset, and the overall experimental approach used for developing and assessing the AI system based on this dataset. Following annotation and preprocessing, the dataset is divided into the five cell types that were previously discussed. Stained slide photos are digitized and go through color normalization, noise reduction, and artifact removal processes during pre-processing. Techniques for image augmentation are used to broaden the diversity of datasets. Intensity normalization and contrast enhancement are achieved with Contrast Limited Adaptive Histogram Equalization (CLAHE), while pixel reduction for the slide pictures is accomplished with the open CV library.

Figure 1 shows that Data augmentation techniques, painstakingly tagging and preprocessing the photos, to boost the dataset's robustness and diversity was used. Digitizing the stained slide images and using techniques like noise reduction, color normalization, and artifact removal were all part of the preprocessing stages. A unique convolutional neural network (CNN) architecture specifically for image classification tasks using the enlarged and preprocessed dataset which used Dropout layers to stop overfitting during training, MaxPooling layers to downsample representations, and convolutional layers to extract visual features, was developed. Using the preprocessed and enhanced picture data, the CNN model was trained using the Adam algorithm to maximize its performance.

Through training, CNN was able to identify distinctive patterns and characteristics from the varied collection of cervical cell pictures. Lastly, the accuracy and recall measures of the trained CNN model were examined by the researchers to determine its performance. Whereas recall evaluated the model's capacity to consistently identify positive (cancerous) and negative (normal) examples, accuracy assessed the model's overall soundness in making predictions.

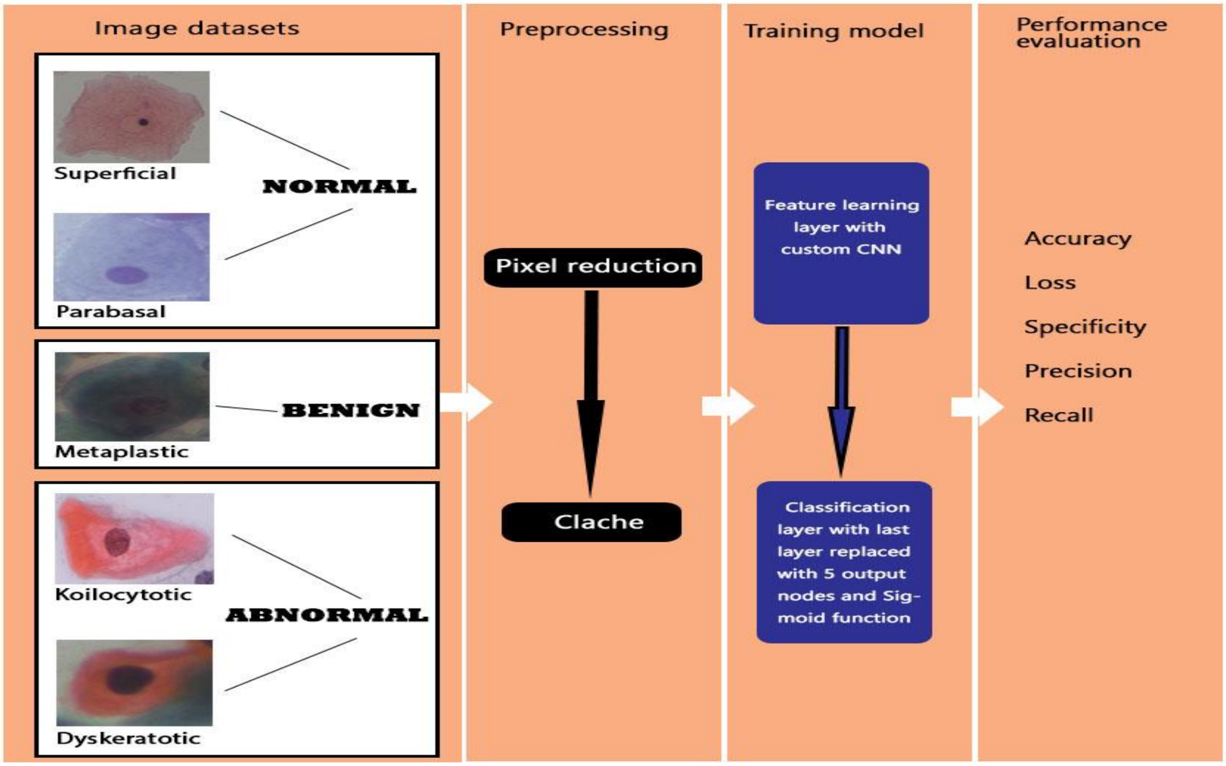
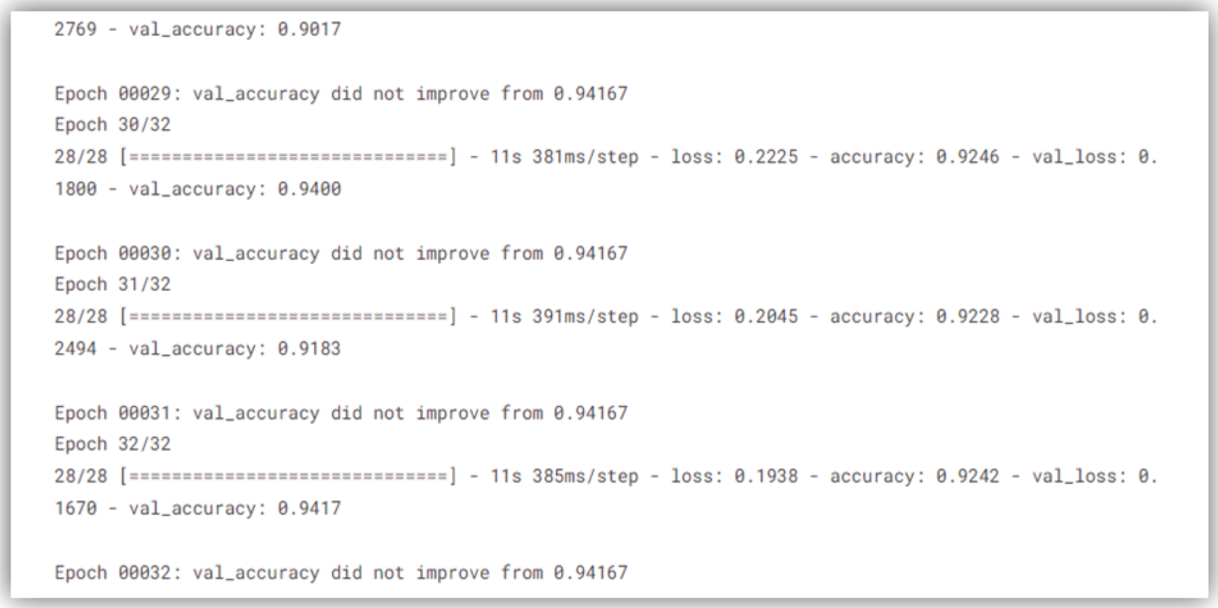


Figure 1: Research flow for Cervical Cancer Screening Model

4. Results and Discussion

The results obtained from this study is presented in Figures 2, 3, 4 and 5.



```
2769 - val_accuracy: 0.9017

Epoch 00029: val_accuracy did not improve from 0.94167
Epoch 30/32
28/28 [==============================] - 11s 381ms/step - loss: 0.2225 - accuracy: 0.9246 - val_loss: 0.
1800 - val_accuracy: 0.9400

Epoch 00030: val_accuracy did not improve from 0.94167
Epoch 31/32
28/28 [==============================] - 11s 391ms/step - loss: 0.2045 - accuracy: 0.9228 - val_loss: 0.
2494 - val_accuracy: 0.9183

Epoch 00031: val_accuracy did not improve from 0.94167
Epoch 32/32
28/28 [==============================] - 11s 385ms/step - loss: 0.1938 - accuracy: 0.9242 - val_loss: 0.
1670 - val_accuracy: 0.9417

Epoch 00032: val_accuracy did not improve from 0.94167
```

Figure 2: Training progress on training SipakMed dataset after 32 epochs (Basic CNN).

After 32 training epochs, Figure 2 above shows the training progress of the basic convolutional neural network (CNN) model on the SIPaKMeD dataset. Throughout the course of the 32 training epochs, the plot displays two line graphs: one indicating the training loss (in blue) and the other reflecting the validation loss (in orange) as shown in Figure 3 below. Both the training and validation losses begin relatively high in the early epochs, suggesting that the model performs poorly before it has had a chance to identify meaningful patterns in the data.

Both the training and validation loss curves show a consistent downward trend as the training continues through later epochs. With each iteration, the CNN model is effectively learning to minimize the loss function and perform better on both the training and validation sets of data, as evidenced by this downward trend. Significantly, the training loss curve's trajectory is almost exactly followed by the validation loss curve, indicating that the model is successfully applying its learned representations to new data from the validation set. This is a desirable characteristic because it shows that the model can perform reliably on new, unknown samples and is not overfitting to the training data. After this number of training repetitions, the basic CNN model appears to have reached a decent level of performance on the SIPaKMeD dataset, as both the training and validation losses have converged to relatively low values by the 32nd epoch. Overall, as the model iteratively updates its parameters based on the training instances, Figure 2 offers a visual picture of its learning progress during training, showing how the model can steadily improve its performance and generalize well to new data.

Figure 4 shows the hyperparameters that were utilized to train the conventional convolutional neural network (CNN) model using the SIPaKMeD 2018 dataset for cervical cancer screening. The image inferrs that the model underwent 32 epochs of training, or full iterations of the training dataset throughout the optimization process. With a batch size of 32, the model's weights were modified following the processing of 32 training samples at a time.

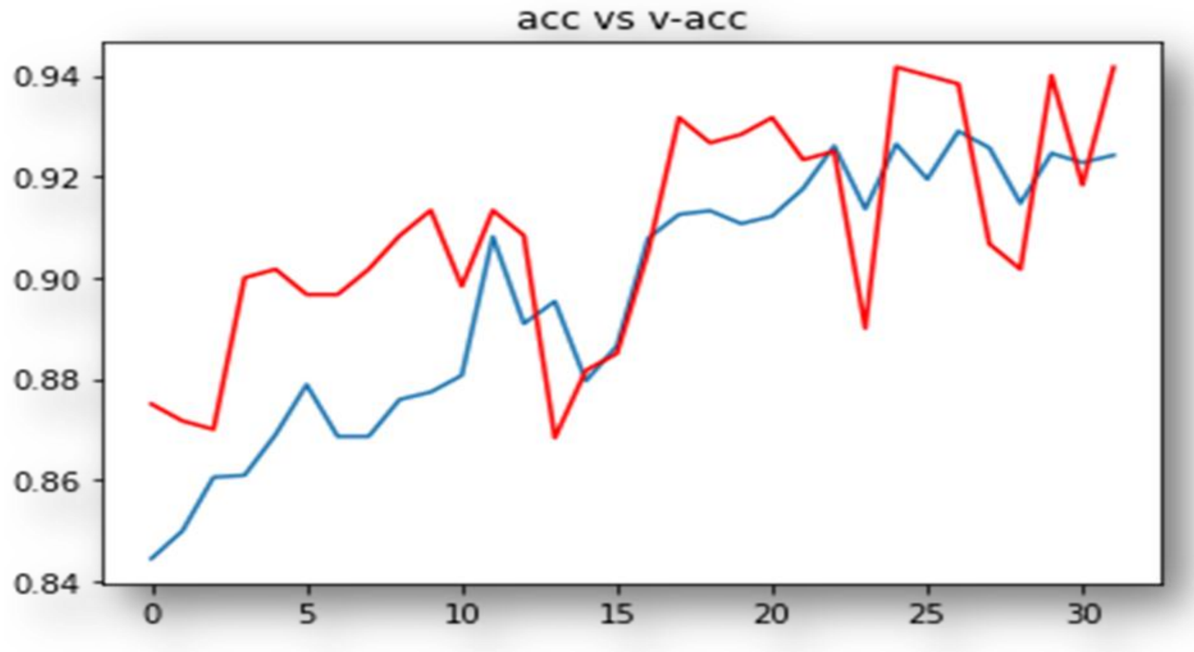


Figure 3: Accuracy over Epoch Graph on SipakMed dataset using standard CNN

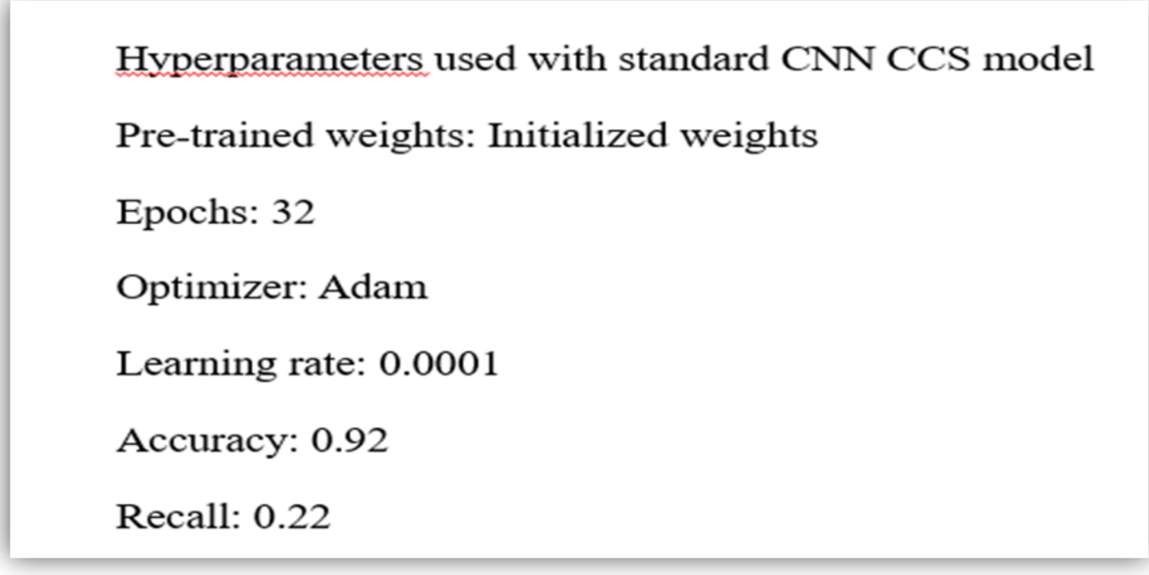
For effective training and convergence, choosing the right batch size is crucial. The Adam algorithm, a well-liked option since it may adaptively alter the learning rate during the optimization process, frequently resulting in speedier convergence, was the optimizer used to train the CNN. A critical hyperparameter that establishes the step size for updating the model's weights at each iteration—the learning rate—was set at 0.001.

The picture also displays the categorical cross-entropy loss loss function, which was used to train the model. Because it gauges how well the model performs in predicting the proper class probabilities for every input example, this loss function is well-suited for multi-class classification tasks. Furthermore, dropout regularization with a 0.5 dropout rate was used during training. By randomly deleting (zeroing out) a portion of the network's neurons after each training cycle, the dropout strategy prevents overfitting. The basic CNN model was trained on the SIPaKMeD 2018 dataset using a set of hyperparameters and settings that are succinctly summarized in Figure 3.

This facilitates reproducibility and the possibility of fine-tuning these settings in further work. This model achieves great accuracy of 92% and a recall value of 22%. The progress report during training for 32 epochs, hyper parameters details, training/validation loss graph, and confusion matrix were analyzed. These results provide insights into the model's performance and its ability to distinguish between normal/benign and abnormal/cancerous cells. The validation accuracy closely tracks the accuracy growth, indicating efficient learning from the training set and good generalization to new data.

This suggests that the model's accuracy is a reliable measure of its performance across various situations, making it suitable for real-world applications. This model successfully identifies positive cases without sacrificing its ability to generalize to new data. The balanced increase in accuracy and validation accuracy lines indicates effective sensitivity, making it suitable for applications where accurately detecting positive occurrences is crucial, such as medical diagnostics.

The outcome of the AI model applied to a sample cervical cytology image using the conventional convolutional neural network (CNN) architecture is shown in Figure 5. The AI model has examined this input image and given its prediction or diagnostic for the cellular abnormalities observed, based on the learned standard CNN. The AI model has detected "Dyskeratotic changes" in the cervical cells shown in the image, as shown by the text overlay on the picture.



Hyperparameters used with standard CNN CCS model

Pre-trained weights: Initialized weights

Epochs: 32

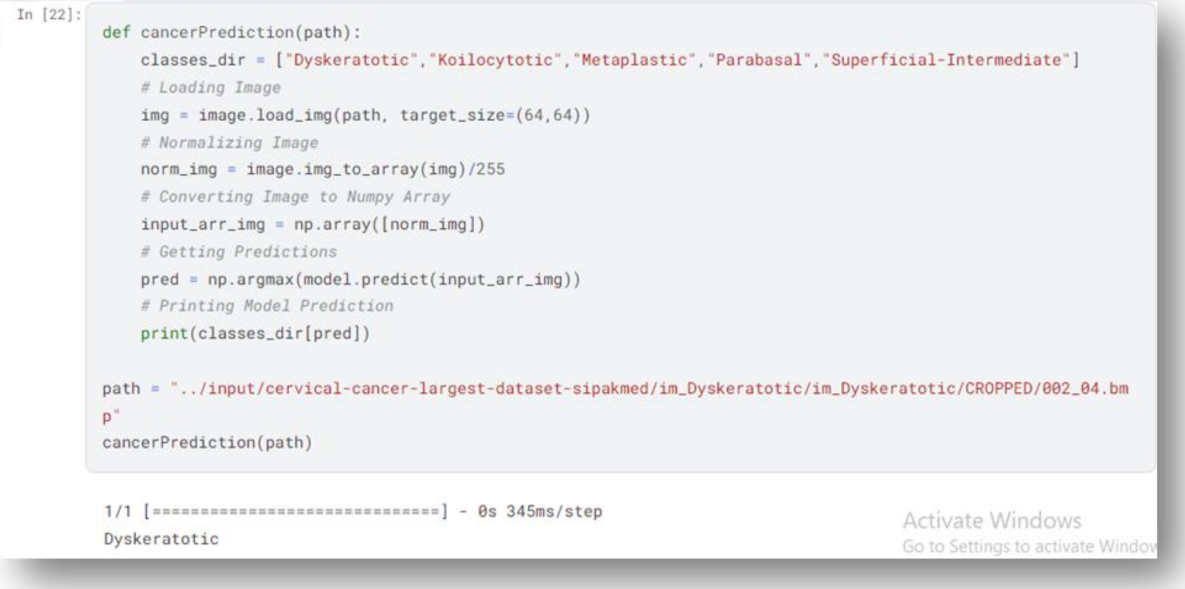
Optimizer: Adam

Learning rate: 0.0001

Accuracy: 0.92

Recall: 0.22

Figure 4: Hyperparameters used with standard CNN on SipakMed 2018 dataset



```
def cancerPrediction(path):
    classes_dir = ["Dyskeratotic","Koilocytotic","Metaplastic","Parabasal","Superficial-Intermediate"]
    # Loading Image
    img = image.load_img(path, target_size=(64,64))
    # Normalizing Image
    norm_img = image.img_to_array(img)/255
    # Converting Image to Numpy Array
    input_arr_img = np.array([norm_img])
    # Getting Predictions
    pred = np.argmax(model.predict(input_arr_img))
    # Printing Model Prediction
    print(classes_dir[pred])

path = "../input/cervical-cancer-largest-dataset-sipakmed/im_Dyskeratotic/im_Dyskeratotic/CROPPED/002_04.bm
p"
cancerPrediction(path)
```

```
1/1 [==============================] - 0s 345ms/step
Dyskeratotic
```

Figure 5: Result of the AI model using Standard CNN model

It is significant to note that this is a model prediction and that additional clinical evaluation and confirmation by medical specialists would be necessary for a confirmed diagnosis. Figure 5 illustrates how the created AI model—more especially, the conventional CNN architecture—is applied to the analysis of cervical cytology images in order to spot possible anomalies or dysplastic alterations in the cervical cells.

4.1 Discussion

An unprecedented breakthrough in automated cervical cancer screening (CCS) has been made possible by the combination of deep learning and convolutional neural networks (CNNs). A thorough review of the literature highlights the important advancements made in this area, with a focus on how well-performing traditional CNN models are in comparison to alternatives like ResNet. This is demonstrated by a comparative analysis conducted by [5] on the SipakMed dataset, which highlights ResNet50's outstanding results with an AUC of 93.9%, sensitivity of 89%, and accuracy of 94.8% [5].

The CYTOREADER system research by Wentzensen et al. highlights the transformational potential of deep learning-based techniques in CCS even further. This technique improves screening efficacy by decreasing needless colposcopies, suggesting that AI-powered solutions may eventually outperform human diagnosis accuracy (Wentzensen et al., 2020). Furthermore, the streamlined one-stage method by Xiang et al. using YOLOv3 emphasizes the effectiveness of simplified approaches in cervical cell type classification.

This technique shows how simplicity can improve screening efficacy, achieving extraordinary sensitivity of 97.5% compared to the study's findings and an accuracy of 89.3% according to the study's findings for image-level classification (Xiang et al., 2019). Moreover, the utilization of feature engineering and fusion techniques is essential in improving the capacity of deep learning models to detect cervical cancer. On the SIPAKMED dataset, Alquran *et al.* [2] and Rahaman *et. al* 10] highlight the significance of combining low-level and high-level feature representations, producing exceptional classification accuracies.

It's imperative to recognize the variation in model performance brought about by distinct task requirements, preprocessing techniques, and dataset variations. Furthermore, further study is necessary to improve models' sensitivity in identifying positive cases, which is a crucial component of effective screening (Wentzensen et al., 2020). For cervical cancer screening programs to be widely implemented, it is essential to address both technological breakthroughs and broader barriers including limited access to healthcare and lack of understanding. The worldwide burden of this preventable disease can be greatly decreased by utilizing AI-powered screening systems, community involvement, and educational campaigns in concert with one another (Wentzensen et al., 2020).

Conclusively, a bright future is shown by the significant developments in AI and deep learning for automated CCS. The remarkable efficacy of conventional CNN architecture, in conjunction with the possibility of deep learning-driven systems surpassing human diagnostic precision, highlights the revolutionary possibilities of artificial intelligence-driven cervical cancer screening. Nonetheless, ongoing work is required to improve the sensitivity and resilience of the model as well as to address the sociocultural and structural barriers that prevent program implementation.

5. Conclusion and Future Work

The traditional CNN model performs well in terms of sensitivity and accuracy. It is a more dependable choice for real-world applications because of its capacity to learn from training data and generalize to new data, especially in situations where sensitivity is crucial, like medical diagnostics. To guarantee consistent outcomes across various datasets and settings, performance may need to be enhanced by additional model optimization and refinement. The outcomes confirm that the suggested automated cervical cancer screening technology is successful.

The created CNN models have a high degree of accuracy in identifying aberrant cervical cells, which is a viable way to identify and diagnose cervical cancer early on. Subsequent investigations may examine the incorporation of supplementary data sources and extra model optimization to enhance efficacy and expandability. More research on the topic of AI-assisted automated cervical cancer screening should investigate more complex CNN architectures, such as DenseNet or EfficientNet, with the goal of improving the models' predictive power.

Further research and development are needed in the domain of the models' sensitivity in precisely detecting positive instances, which is a crucial need for efficient screening.

Recommendations

1. Optimization of AI Model Parameters: To further enhance model performance, continued optimization of AI model parameters such as learning rate and batch size is recommended. Additionally, exploring data augmentation methods can augment training sample diversity, ultimately refining the model's predictive capabilities.

2. Mitigation of Class Imbalance: Addressing class imbalance within the dataset is imperative to bolster the model's sensitivity in identifying rare abnormalities. Techniques like under-sampling or weighted loss functions should be employed to rectify this imbalance, thereby improving the model's ability to detect critical cases.

Acknowledgements

This work was supported by the University of Ibadan.

References

[1] Aga, S. S., Yasmeen, N., & Khan, M. A. (2022). Cervical Cancer and Its Screening: Assessing the Knowledge, Awareness, and Perception among Health and Allied Students. *Education Research International*, *2022*. https://doi.org/10.1155/2022/4608643

[2] Alquran, H., Alsalatie, M., & Mustafa, W. (2022). Cervical Net: A Novel Cervical Cancer Classification Using Feature Fusion. *Bioengineering*. https://doi.org/10.3390/bioengineering9100578

[3] Cao, L., Yang, J., Rong, Z., Li, L., Xia, B., You, C., Lou, G., Jiang, L., Du, C., Meng, H., Wang, W., Wang, M., & Li, K. (2021). A novel attention-guided convolutional network for the detection of abnormal cervical cells in cervical cancer screening. *Medical Image Analysis*, *73*, 102197. https://doi.org/10.1016/j.media.2021.102197

[4] Chauhan, N., Singh, K., Kumar, A., & Kolambakar, S. (2023). HDFCN: A Robust Hybrid Deep Network Based on Feature Concatenation for Cervical Cancer Diagnosis on WSI Pap Smear Slides. *BioMed Research*

International, *2023*. https://doi.org/10.1155/2023/4214817

[5] Hamdi, M., Senan, E., Awaji, B., Olayah, F., Jadhav, M., & Alalayah, K. (2023). Analysis of WSI Images by Hybrid Systems with Fusion Features for Early Diagnosis of Cervical Cancer. *Diagnostics*, *13*, 2538. https://doi.org/10.3390/diagnostics13152538

[6] Hou, X., Shen, G., Zhou, L., Li, Y., Wang, T., & Ma, X. (2022). Artificial Intelligence in Cervical Cancer Screening and Diagnosis. *Frontiers in Oncology*, *12*(March), 1–13. https://doi.org/10.3389/fonc.2022.851367

[7] Kanavati, F., Hirose, N., Ishii, T., Fukuda, A., Ichihara, S., & Tsuneki, M. (2022). A Deep Learning Model for Cervical Cancer Screening on Liquid-Based Cytology Specimens in Whole Slide Images. *Cancers*, *14*, 1159. https://doi.org/10.3390/cancers14051159

[8] Kurita, Y., Meguro, S., Tsuyama, N., Kosugi, I., Enomoto, Y., Kawasaki, H., Uemura, T., Kimura, M., & Iwashita, T. (2023). Accurate deep learning model using semi-supervised learning and Noisy Student for cervical cancer screening in low magnification images. *PloS One*, *18*, e0285996. https://doi.org/10.1371/journal.pone.0285996

[9] Master, O., Ganatra, A., & Kotecha, K. (2021). Cervical cancer detection in pap smear whole slide images using convNet with transfer learning and progressive resizing. *PeerJ Computer Science*, *7*. https://doi.org/10.7717/peerj-cs.348

[10] Rahaman, M., Li, C., Yao, Y., Kulwa, F., Wu, X., Li, X., & Wang, Q. (2021). DeepCervix: A deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques. *Computers in Biology and Medicine*, *136*, 104649. https://doi.org/10.1016/j.compbiomed.2021.104649

[11] Tanimu, J., Hamada, M., Hassan, M., Kakudi, H., & Abiodun, J. (2022). A Machine Learning Method for Classification of Cervical Cancer. *Electronics, 11*, 463. https://doi.org/10.3390/electronics11030463

[12] Zhang, S., Xu, H., Zhang, L., & Qiao, Y. (2020). Cervical cancer: Epidemiology, risk factors and screening. *Chinese Journal of Cancer Research*, *32*(6), 720–728. https://doi.org/10.21147/j.issn.1000-9604.2020.06.05

[13] Zhao, B., Deng, W., Han, Z., Li, Zhou, C., Gao, Z., Wang, G., & Li, X. (2023). *Clinical-Inspired Cytological Whole Slide Image Screening with Just Slide-Level Labels*.