**Title: Design and Implementation of a Custom Computer Vision Model Using Classical Deep Learning Techniques with Explainable AI for Image Classification**

**Abstract**

The abstract should contain the following parts

Backgroung and introduction of the problem

Your proposed methodology describing what you have done

Brief details of the results obtained

Brief conclustion that state the usefulness of your work.

**Keywords:** keyword 1; **:** keyword 1; **:** keyword 1; **:** keyword 1

# **Introduction**

Expand on the nature of the problem you are trying to solve. Mention the traditional methods that have been used and the different AI techniques that have been adopted to solve the problem. State the limitations of these AI techniques previously used and how you inten to overcome the limitations to solve the proble. Highlight the significant contribution of your work. An example is provided below on how to write the contribution of your work:

The work presented here has resulted in the following significant contributions:

* The introduction of SegmentNet, a new AI model specifically designed for medical image segmentation. This model is built upon the UNet architecture, thus, the UNet is reconstructed to fit into the functionality of the SegmentNet. incorporating both our proposed distance-aware mechanisms and a local feature extraction block to the U-structured network.
* This report proposes a distance-aware mechanism that focuses on the non-related features in the input image. This helps the model to learn well and become more robust in handling varying medical image modalities.

Also, an example has been provided below to guide you on how to conclude the introduction section by stating how the remaining sections of your report will be structured:

**The structure of the remaining sections of this report is as follows. Section 2 provides a brief review of relevant studies. In Section 3, we discuss our proposed framework including the datasets used. In Section 4, we present the results and discussion of our experiments, showcasing the strong performance of our approach on two datasets and in comparison with other state-of-the-art models. Finally, in Section 5, we conclude our findings and outline potential avenues for future research.**

# **Related work**

You should report relevant and recent literature related to the problem your are addressing. Mention what the literature did, what method they used, the result they obtained, and the their limitations. An Example is provided below:

Recent advancements in the UNet architecture have been primarily aimed at enhancing the performance of UNet models. These improvements have been achieved through two main approaches: uniform scaling of the network and leveraging pre-trained CNN models trained on the ImageNet dataset as encoders. Zhou et al. [11] propose two notable variants of the UNet architecture. The first variant is Wide UNet, which achieves uniform scaling by increasing the number of filters in both the encoder and decoder subnetworks of UNet.

NOTE: Make sure tocorrectly reference every literature you have included in your report. Provide references to support your claims.

# **3. Material and Method**

## **3.1 Dataset**

Mention the dataset collection process, the database repository, the nature of the dataset and other information. An example is provided below:

To build and evaluate the proposed SegmNet, we use the diabetic retinal funds (DRF). This is to show the capability of our model to segment various medical images and provide a generalization solution for the segmentation task. In this section, we introduce the target medical datasets.

1. ***Diabetic Retinal Fundus (DRF):*** The images of this dataset are specialized retinal images that provide detailed and high-resolution visualization of the human eye's fundus region. At the time of harnessing the data, the database provided by [3] comprises a total of 15 images depicting healthy patients, 15 images depicting patients with diabetic retinopathy, and 15 images depicting patients with glaucoma. Some samples taken at random from this dataset alongside their masks are shown in Figure

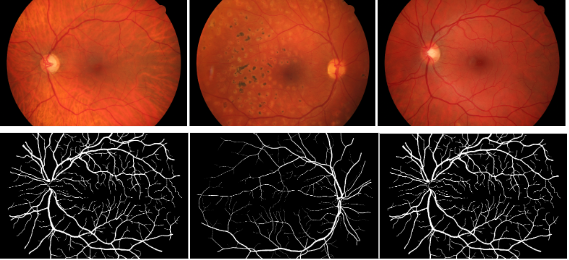


Figure 1. Random Samples Of The Diabetic Retinal Fundus Images And Their Masks

## **3.2 Proposed Model**

Describe your own proposed medel. An example is provided below:

The SegmentNet like most other models used for segmentation is parted into two – the encoder part and the decoder part with the distance-aware mechanism, skip connection, and local feature interaction extraction block. Figure 2 shows the detailed deep learning architecture of the proposed SegmentNet. The encoder portion of the network comprises five convolution levels from the upper part of it to the bottom part.

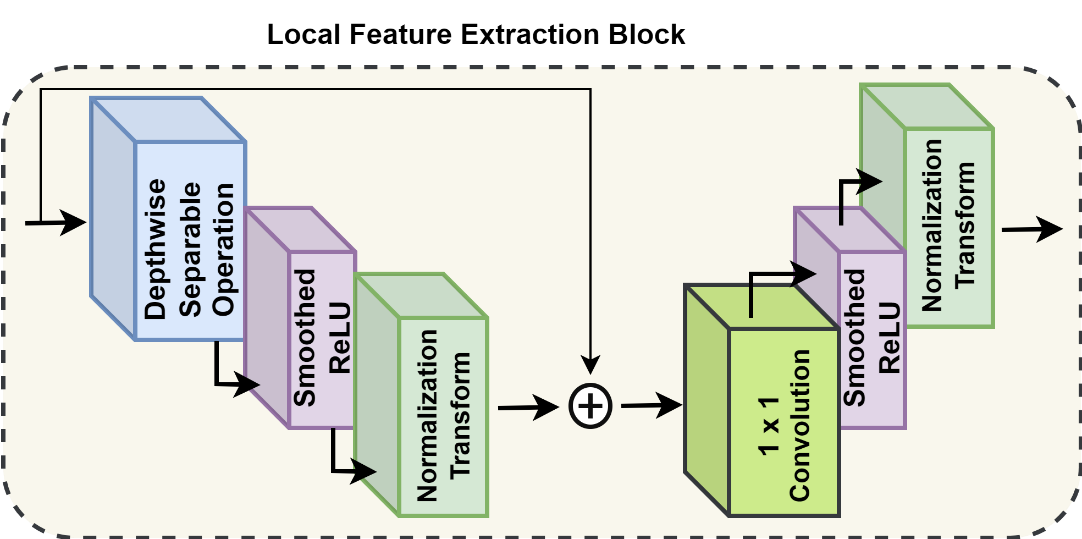


Figure 2**.** The Architectural Overview of the Local Feature Extraction Block

One layer of the local feature extraction block comprises a depthwise separable operation whose kernel size is k x k and a 1 x1 convolution. The composition of the channels of the input feature map is equivalent to the number of groups in the depthwise separable operation kernel. Every channel is also equipped with a Smoothed ReLU activation alongside a Normalization Transform. The architectural design of this block is shown in Figure 4. The normalization transform is expressed by the following equations (1) and (2)

(1)

(2)

Where *NT* stands for Normalization Transform, stands for Smoothed ReLU activation, and is the feature map got as output from layer *l* in the block of the local feature extraction block. The layers of this block module produce feature maps that have the capacity to retain the same size and resolution before the extracted features from it are upsampled directly.

## **3.3 Evaluation Strategy**

State the different evaluation metrics you have used in your work and their formulars (equations)

An Example has been provided below:

These metrics used in this report include Recall or Sensitivity, Jaccard Index or Intersection over Union (IoU), F1-value, Specificity, and Accuracy. Each of the metrics is mathematically expressed as follows:

|  |  |
| --- | --- |
|  | (3) |
|  | (4) |
| * Sensitivity/ | (5) |
|  | (5) |
|  | (7) |
|  | (8) |

This metric is related to the dice score also called the Dice-Sørensen coefficient. The Jaccard index otherwise called IoU assesses the overlap or intersection between the ground truth and the predicted segmentation as a ratio of the union of the predicted segmentation and the ground truth [1,3].

## **3.4. Environment Execution**

Mention the environmental setup of the hardware(Computer) and software you have used to execute your code. An example has been provided below:  
The models discussed in this study were implemented using Python in a Lenovo Windows system with 64 GB of RAM. The system is equipped with an Nvidia GeForce RTX 3080 chip, known for its high-performance capabilities in handling intensive computational tasks. Additionally, the system features an Intel® Core(TM) i9-10850K CPU with a base clock speed of 3.60 GHz and a maximum turbo frequency of 5.20 GHz, providing substantial processing power for the model training and evaluation processes.

# **4. Experimental Results**

Provide the experimental result of your work. An example has been provided below:

This section presents the segmentation results using three datasets separately: (1) Breast ultrasound images (BUSI), (2) Chest X-ray images (CXRI), (3) and Diabetic retinal Fundus images (DRFI). For each dataset, the segmentation performance of the proposed SegmentNet is directly compared with seven state-of-the-art AI models: UNet [4], AttentionUNet [34], DeepUNet [35], UNetplus [5], UNet2plus [6], UNet3plus [4], and CMUNet [6]. The proposed SegmentNet is directly compared with the other AI models using the same training settings, environment, and dataset. Comparing the proposed SegmentNet with other recent models establishes a benchmark performance in medical image segmentation. Understanding the comparative performance highlights the unique contributions and advantages of our model.

**4.1. Performance Results using the dataset**

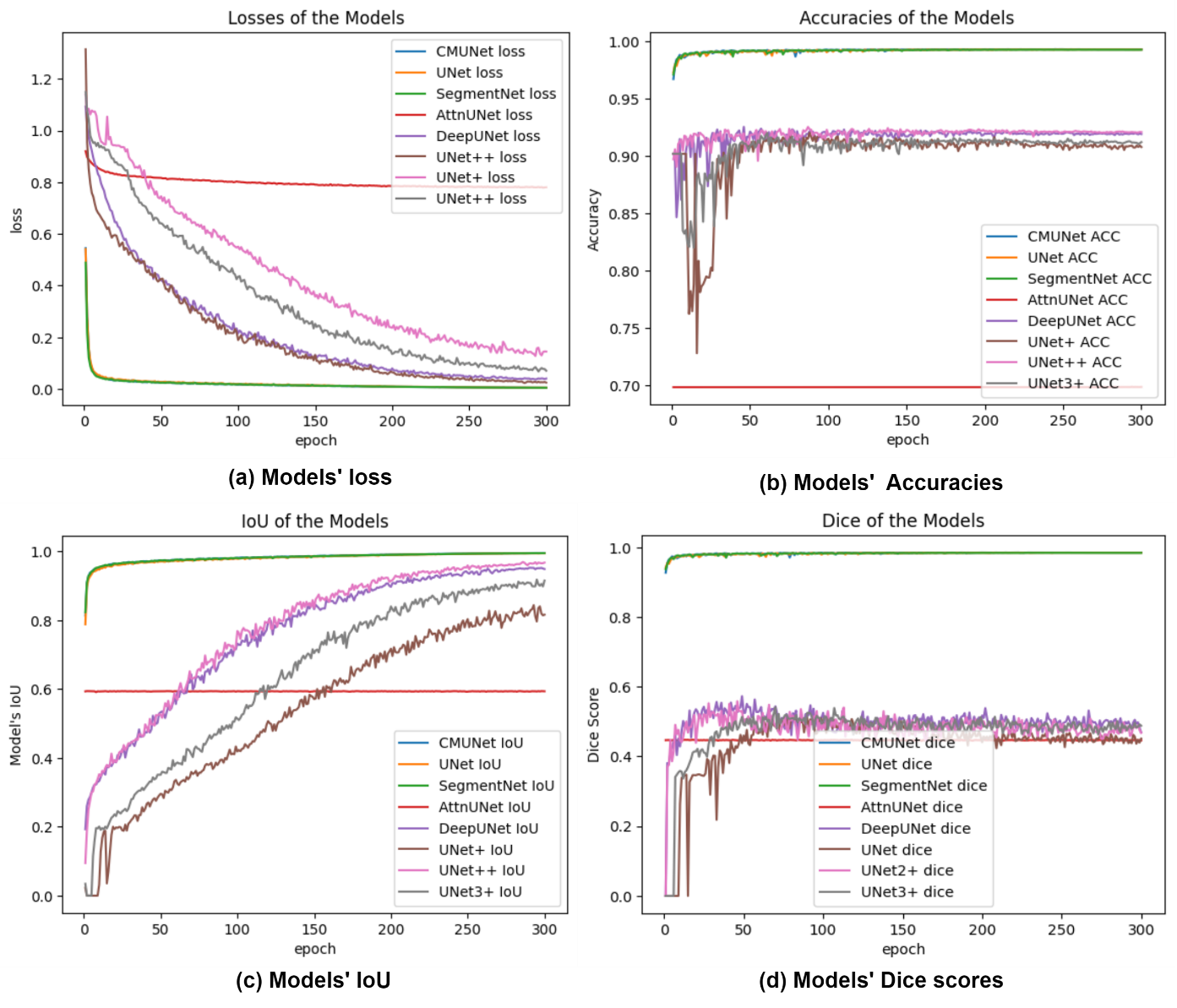
Provide the performance result of your work. An example has been provided below:

In this section, we present the segmentation results of the proposed AI model against the other seven AI models. We perform the segmentation task twice w.r.t the number of classes in the BUSI dataset. The first experiment is conducted over the benign and malignant classes and the evaluation results are shown in **Table 1**.

**Table 1**. Evaluation results using the datset over two classes: benign and malignant.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **AI Model** | **IoU** | **Accuracy** | **SP** | **F1** | **Dice** | **SE** |
| UNet | 93.98 | 92.53 | 99.89 | 57.57 | 57.57 | 63.32 |
| AttentionUNet | 69.81 | 69.84 | 56.05 | 47.80 | 47.80 | 87.31 |
| DeepUNet | 95.23 | 92.53 | 100 | 57.30 | 57.30 | 71.34 |
| UNetplus | 84.37 | 92.03 | 100 | 51.78 | 51.78 | 73.03 |
| UNet2plus | 96.91 | 92.53 | 99.47 | 55.58 | 55.85 | 59.64 |
| UNet3plus | 91.81 | 92.01 | 100 | 54.13 | 54.13 | 60.92 |
| CMUNet | 96.93 | 92.12 | 99.73 | 55.54 | 55.54 | 73.16 |
| SegmentNet | 96.96 | 92.12 | 100 | 64.79 | 64.79 | 73.59 |

To further show the performance of the models on predicting the cancer tumors in the breast, we have shown in Figure 2. the significant differences in the predicted masks by the models from the various networks implemented on the dataset. Our proposed model had the capacity to locate all the tumors in each of the images and at the corresponding positions and in comparison to other models’ performances, it outperformed them.



## **4.2. Discussion**

Provide in-depth discussion of your work. An example has been provided below:

Looking at the results for the different models on the BUSI dataset in Table 1 and Table 2 UNet, DeepUNet, UNet+, UNet2+, CMUNet, and SegmentNet demonstrate relatively high IoU, Accuracy, and SP scores, indicating their effectiveness in accurately segmenting the desired objects. AttentionUNet and UNet3+ show notably lower performance with very low IoU, Accuracy, SP, and F1 scores. This suggests that these models may struggle to accurately capture the object boundaries and discriminate between the target and background. SegmentNet achieves the highest performance across most metrics, with the highest IoU, Accuracy, SP, F1, Dice, and SE scores. This indicates that SegmentNet performs exceptionally well in segmenting the desired objects and achieves a high level of accuracy. SegmentNet stands out as the top-performing model in terms of accuracy and segmentation quality, while AttentionUNet and UNet3+ exhibit weaker performance. The segmentation results from this is dataset shown in Figure 6. Based on the results in the afore-discussed Tables, SegmentNet demonstrates superior performance in accurately segmenting breast ultrasound images across multiple classes (benign, malignant, and normal). Its high IoU, F1 score, Dice score, sensitivity, specificity, and accuracy highlight its capability to accurately identify and delineate regions of interest, making it a reliable choice for medical imaging applications.

## **4.3 Fair comparison with othe Deep Learning Models**

Provide fair comparison of your work with other pre-trained deep learning models. An example has been provided below:

Some of the work that has been presented in the literature on the BUSI dataset was carried out using 2 classes of the dataset, in [38], Tang et al used only the benign and the malignant component of the dataset since the masks of the normal subset have nothing in them. With their network CMUNet, the result obtained was compared to those from our proposed SegmentNet model. Whereas for the two classes Tang et al [38] obtained an IoU of 73.27 ± 0.43, our model obtained 96.96% which is better although when their method was applied by us an IoU of 96.93 was obtained. Accuracy-wise, their model had the same performance as ours (92.12) with our experimental setup, however, they recorded 97.33 ± 0.14%. Using the same architecture as they, we obtained better f1 scores, dice, specificity, and sensitivity although what was reported in the report differs from our obtained result using their approach. Ru et al [39] also used 2 classes of the dataset for their experiment obtaining 77.76% Dice, 69.91% IoU, 77. 65% precision, 84.28% sensitivity, and 98.47% specificity all measured in percentage. The results they obtained however trail the ones yielded by our model in IoU and specificity while outperforming ours slightly in other metrics.

## **4.4 Comparison with Existing Literature**

Provide direct comparison of your work with other prior works. An example has been provided below:

Some of the work that has been presented in the literature on the BUSI dataset was carried out using 2 classes of the dataset, in [38], Tang et al used only the benign and the malignant component of the dataset since the masks of the normal subset have nothing in them. With their network CMUNet, the result obtained was compared to those from our proposed SegmentNet model. Whereas for the two classes Tang et al [38] obtained an IoU of 73.27 ± 0.43, our model obtained 96.96% which is better although when their method was applied by us an IoU of 96.93 was obtained. Accuracy-wise, their model had the same performance as ours (92.12) with our experimental setup, however, they recorded 97.33 ± 0.14%. Using the same architecture as they, we obtained better f1 scores, dice, specificity, and sensitivity although what was reported in the report differs from our obtained result using their approach. Ru et al [39] also used 2 classes of the dataset for their experiment obtaining 77.76% Dice, 69.91% IoU, 77. 65% precision, 84.28% sensitivity, and 98.47% specificity all measured in percentage. The results they obtained however trail the ones yielded by our model in IoU and specificity while outperforming ours slightly in other metrics.

# **5. Conclusion, Limitation, Future work**

Provide a detailed conclusion, limitation and future of your work. An example has been provided below:

In this report, we presented SegmentNet, a novel architecture for binary and multi-class segmentation tasks with high performance on different image modalities. SegmentNet is an encoder-decoder network that integrates a local feature extraction mechanism that focuses on the local feature interaction and the multi-focus distance-aware mechanism which focuses on other features in the well-known UNet segmentation architecture to achieve accurate and precise segmentation results. The major idea behind this work is to leverage the strength of local feature extraction and distance-aware mechanisms in enhancing the performance of segmentation tasks. Our research is anticipated to serve as a supplementary tool for medical professionals and offer insights into the interpretability of breast tumor and chest X-ray segmentation tasks. In the future, we intend to investigate the most effective methods for bridging the model into a lighter-weight form since it is robust and can generalize well..

**References**

Provide relevant and current references using IEEE referencing style. An example has been provided below:

[1] Monday HN, Li J, Nneji GU, Nahar S, Hossin MA, Jackson J, et al. COVID-19 Diagnosis from Chest X-ray Images Using a Robust Multi-Resolution Analysis Siamese Neural Network with Super-Resolution Convolutional Neural Network. Diagnostics 2022;12. https://doi.org/10.3390/DIAGNOSTICS12030741.

[2] Liu W, Luo J, Yang Y, Wang W, Deng J, Yu L. Automatic lung segmentation in chest X-ray images using improved U-Net. Sci Reports 2022 121 2022;12:1–10. https://doi.org/10.1038/s41598-022-12743-y.

[3] Singh VK, Abdel-Nasser M, Akram F, Rashwan HA, Sarker MMK, Pandey N, et al. Breast tumor segmentation in ultrasound images using contextual-information-aware deep adversarial learning framework. Expert Syst Appl 2020. https://doi.org/10.1016/j.eswa.2020.113870.

[4] Ejiyi CJ, Qin Z, Adetunji SA, Happy MN, Nneji GU, Ukwuoma CC, et al. Comparative Analysis of Building Insurance Prediction Using Some Machine Learning Algorithms. Int J Interact Multimed Artif Intell 2022;7:75–85. https://doi.org/10.9781/ijimai.2022.02.005.

[5] Guleria K, Sharma S, Kumar S, Tiwari S. Early prediction of hypothyroidism and multiclass classification using predictive machine learning and deep learning. Meas Sensors 2022;24:100482. https://doi.org/10.1016/J.MEASEN.2022.100482.

[6] Agrawal T, Choudhary P. Segmentation and classification on chest radiography: a systematic survey. Vis Comput 2021 393 2022;39:875–913. https://doi.org/10.1007/S00371-021-02352-7.