An Improvised Deep Learning Model with Loss functions Fusion for Image Classification

Dr. Geetha Ramani[1], Kavyadarshnee K S[2], Sreenithi B[2], Devnath R[2] and SethuVignesh V[2]

[1] Professor, Department of Information Technology, Anna University (MIT Campus)
[2] B.Tech, Department of Information Technology, Anna University (MIT Campus)

[1]rgeetha@yahoo.com, [2]darshnee29@gmail.com, [3]srinithisriram1@gmail.com

***Abstract*—:** Deep learning has transformed image classification by enabling models like CNNs to automatically learn features from data, overcoming the limitations of traditional manual feature extraction methods. The VGG architecture, known for its simplicity and effectiveness, is widely used in various image classification tasks.

This project aims to enhance image classification accuracy by leveraging a fusion of multiple loss functions using U-Net, VGG16, and CNN models across datasets like the DeepGlobe Land Cover Dataset, BraTS Brain Tumor Dataset, Tomato Leaf Disease Dataset, and Crack Segmentation Dataset. This project comprises preprocessing of images, training models with the fusion of loss functions, and evaluating their performance against key metrics. A novel approach is proposed to combine loss functions such as Cross-Entropy, Focal Loss, Tversky Loss, Dice Loss, and Smooth L1 Loss to address challenges in convergence and performance, especially with complex and imbalanced datasets. Different activation functions and n-gram combinations were explored to further enhance the model's ability to generalize and learn from the data. These techniques aim to create a more robust and versatile model capable of handling image classification tasks.

The development of this project was carried out using Python as the primary programming language, with extensive use of frameworks such as TensorFlow PyTorch for model implementation and training. Google Colab served as the primary platform for experimentation and model training, leveraging its GPU support for efficient computations.

Index Terms*—Dice Loss, Loss Function Fusion, Image Classification, Deep Learning.*

I. INTRODUCTION

Deep learning models have revolutionized image classification, making it possible for machines to interpret visual data with unprecedented accuracy. Unlike traditional methods that rely on manual feature extraction, deep learning models automatically learn hierarchical features directly from raw data. Convolutional Neural Networks (CNNs), a cornerstone of deep learning, have demonstrated exceptional performance in identifying complex patterns and structures in images. Models namely VGG, ResNet, and U-Net have become the backbone of various applications, from medical imaging and satellite image analysis to object detection and autonomous driving. Their ability to process large-scale datasets and adapt to diverse challenges has made them indispensable in modern computer vision tasks.

The fusion of loss functions plays a crucial role in enhancing the performance of deep learning models for image classification. Tversky Loss generalizes Dice Loss by incorporating parameters that control sensitivity and specificity, making it especially useful for tasks like medical image segmentation where precision and recall need to be balanced. Dice Loss, commonly used in segmentation tasks, measures the overlap between predicted and true class labels, helping improve accuracy, particularly in imbalanced datasets. Focal Loss addresses class imbalance by focusing more on hard to-classify, misclassified samples and down weighting easy examples, which improves learning from difficult cases. Smooth L1 Loss, a combination of L1 and L2 loss, reduces sensitivity to outliers and provides smoother gradients for optimization, making it effective for object detection and regression tasks. Categorical Cross-Entropy Loss is used in multi-class classification to measure the difference between the predicted probability distribution and the true class label, guiding the model to correctly classify each category. Binary Cross-Entropy Loss is suited for binary classification tasks and 2 helps the model distinguish between two classes by calculating the difference between predicted probabilities and true binary labels. Lastly, Mean Squared Error (MSE) Loss is used for regression tasks to measure the average squared difference between predicted and actual values, but can also be applied in classification tasks where class probabilities are treated as continuous. By leveraging advanced loss functions, activation mechanisms, and optimization techniques, these models continue to push the boundaries of image classification, transforming industries and enabling innovative solutions.

II. LITERATURE SURVEY

Recent advancements in deep learning have significantly enhanced image classification tasks, particularly in medical image segmentation. Various architectures and innovative loss functions have been proposed to improve model performance. Ocal et al. [1] introduced the DSBAV-Net, a V-shaped network designed for medical image segmentation, while Liu et al. [2] presented the SMRU-Net, which utilizes channel-space separate attention and depthwise separable convolutions for skin disease image segmentation. Similarly, Zhang [6] applied a U-Net architecture to enhance precipitation forecast accuracy, and Rawat [8] proposed the 3D U-Net-Norm for better generalization of BraTS images. These architectural innovations demonstrate the flexibility of deep learning models in addressing diverse image classification challenges across different domains.

The fusion of loss functions has emerged as a key strategy for improving model accuracy and generalization across complex tasks. Researchers have developed various fusion approaches, combining multiple techniques and loss functions to optimize model performance. Bougourzi et al. [15] proposed a Transformer-CNN architecture combined with an edge loss function for multi-class pneumonia infection segmentation, while Ke et al. [12] utilized adaptive Tversky loss with a scale-aware UNet++ model for better accuracy. Chen et al. [9] introduced confidence-weighted mutual supervision for cross-modality image segmentation, and Ming et al. [10] demonstrated the benefits of gradient-optimized Dice loss for 5 addressing challenging segmentation tasks. These strategies highlight the power of integrating complementary loss functions to tackle diverse and complex problems in image classification.

III. SYSTEM ARCHITECTURE AND DESIGN

This chapter delves into the architectural design of the proposed image classification system. The system leverages the VGG architecture as the backbone, combined with a robust preprocessing pipeline and a strategic selection of loss functions. The system is designed to efficiently handle image classification tasks, particularly those involving complex datasets and challenging scenarios.

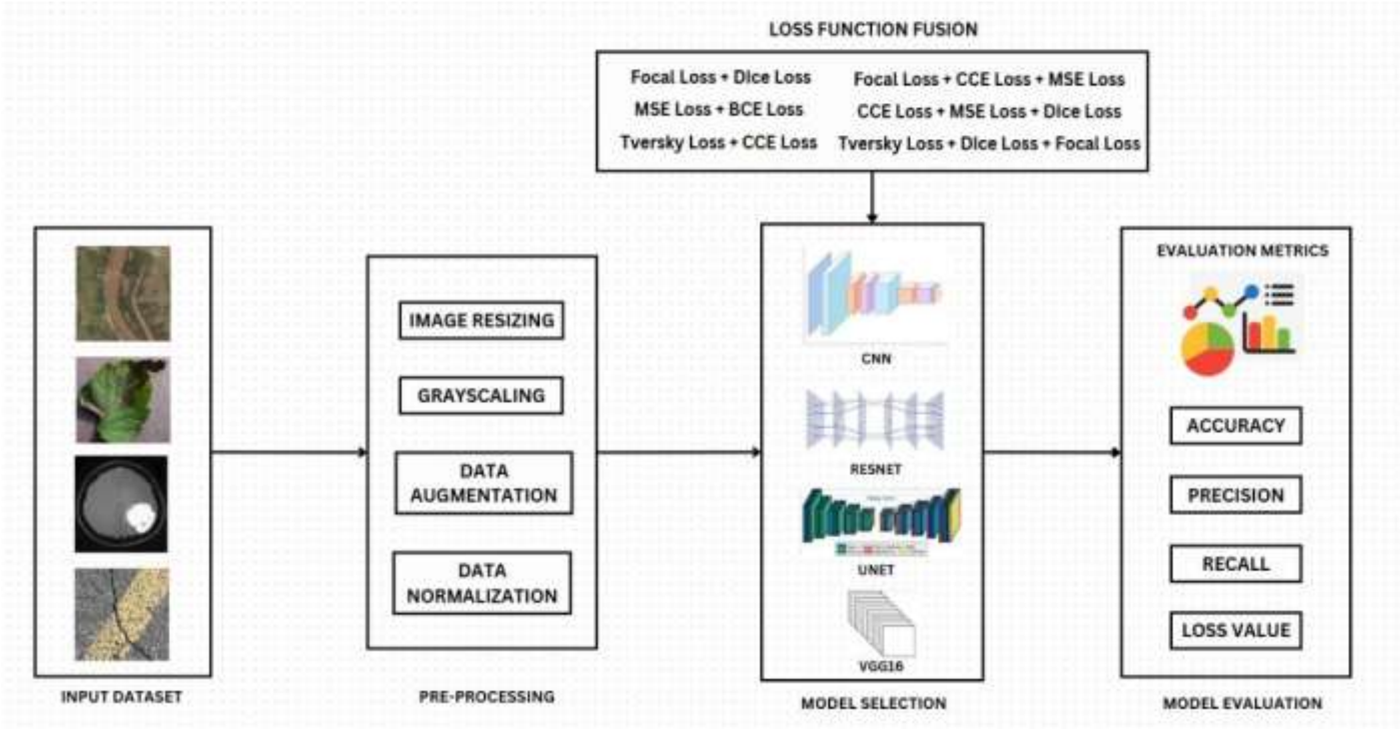

Figure 1: Proposed Model Architecture

Dataset Description

This approach leverages four diverse datasets to comprehensively evaluate the performance of image classification models**:**

- DeepGlobe Land Cover Dataset: This dataset presents a significant challenge due to the high variability and complexity of land cover types. It contains satellite images with seven distinct classes, including urban land, agriculture land, rangeland, forest land, water bodies, barren land, and unknown land. Count on each label in present in Table 3.1. Accurate classification of these classes is crucial for various applications, such as urban planning and environmental monitoring.

Table 1 : DeepGlobe Land Cover dataset description

Class labels	Count
Urban land	800
Agricultural land	722
Range Land	521
Forest Land	191
Water	476
Barren	422
Unknown	28

- Brain Tumor Dataset: This dataset comprises medical images containing brain scans with four tumor categories: glioma, meningioma, notumor, and pituitary tumors as mentioned in Table 3.2. Accurate segmentation of tumor regions is critical for diagnosis and treatment planning. This dataset presents challenges due to the subtle visual differences between tumor types and the presence of healthy tissue.

Table 2: Brain Tumor dataset description

Class labels	Count
Glioma	300
pituitarytumors	300
meningioma	306

- Tomato Leaf Disease Dataset: This dataset consists of images of tomato leaves with five disease classes, including bacterial, early blight, healthy, late blight, and leaf mold. Count on each label is present in Table 3.3. Accurate classification of these diseases is crucial for early detection and disease management.

Table 3: Tomato Leaf dataset description

Class labels	Count
Tomato mosaic virus	1000
Target spot	1000
Bacterial spot	1000
Late blight	1000
Tomato Healthy	1000

- Crack Segmentation Dataset: This dataset contains images with cracks and other features, requiring accurate segmentation of fine details. This dataset described in Table 3.4 presents challenges due to the variability in crack patterns, lighting conditions, and image quality. Accurate crack detection is crucial for infrastructure maintenance and safety inspections.

Table 4: Crack segmentation dataset description

Class labels	Count
CFD	100
Volker	878
Cracktree	175
Sylvie Chambon	157

By utilizing these diverse datasets, the project aims to evaluate the model's ability to handle various image classification and segmentation tasks, including multi-class classification, binary classification.

Proposed Modules

1. Data Preprocessing

Data preprocessing is a critical step in any machine learning pipeline, and image classification is no exception. The following preprocessing steps are employed to ensure optimal model performance:

- Image Loading and Resizing: Images are loaded from the dataset and resized to a consistent size (e.g., 224x224 pixels) to match the input requirements of the VGG model.
- Gray scaling: Images are converted to grayscale to reduce dimensionality and computational complexity while potentially improving performance in certain scenarios.
- Data Augmentation: To improve the model's generalization ability, data augmentation techniques such as random cropping, flipping, rotation, and color jittering are applied to artificially increase the size and diversity of the training dataset.
- Data Normalization: Image pixel values are normalized to a specific range (e.g., 0 to 1) to enhance the model's convergence and stability during training. By implementing these preprocessing steps, the input data is pre-pared effectively for the subsequent stages of the image classification pipe-line.

2. Loss Function Fusion

The fusion of loss functions has been successfully applied across multiple datasets to optimize the performance of deep learning models, particularly in image classification tasks. In the DeepGlobe Land Cover Dataset, Tversky Loss was used to address class imbalances between different land cover types, such as barren and unknown land, while Dice Loss enhanced segmentation accuracy by improving the overlap between predicted and true land boundaries. Similarly, in the Brain Tumor Dataset, Focal Loss was employed to prioritize hard-to-classify tumor regions, and Dice Loss improved the 10 accuracy of tumor segmentation, ensuring precise localization of small tumors.

The fusion of these loss functions ensured that the model could effectively handle the complex task of brain tumor classification and segmentation. Additionally, for the Tomato Leaf Disease Dataset, Tversky Loss addressed class imbalances between disease types like bacterial and early blight, while Smooth L1 Loss im-proved the model's ability to detect fine-grained features in the tomato leaves, especially under varying lighting conditions. Finally, in the Crack Segmentation Dataset, Focal Loss was used to focus on difficult-to-detect cracks, and Dice Loss ensured better segmentation of crack regions, improving the model's ability to detect small and subtle crack patterns in infrastructure images. After applying this fusion of loss functions, the model undergoes training, where it learns to optimize for these various objectives, enhancing its ability to generalize across a range of image classification and segmentation tasks. By combining these loss functions, the model is encouraged to learn from different aspects of the data, leading to im-proved generalization and robustness.

3. Model Selection and Training

The VGG model, a deep convolutional neural network, is employed as the back-bone for the image classification task. The architecture comprises the following key components:

- Convolutional Layers: Multiple convolutional layers extract features from the input images at different levels of abstraction.

- Pooling Layers: Pooling layers, such as max pooling, reduce the spatial dimensions of the feature maps, thereby reducing computational complexity and preventing overfitting.
- Fusion of Loss Functions: Combines multiple loss functions (e.g., adaptive Tversky loss, Dice loss) during training to optimize accuracy, handle class imbalances, and improve boundary detection. 11
- Fully Connected Layers: Fully connected layers combine the extracted features from the convolutional layers to make the final classification.
- Softmax Layer: The softmax layer outputs the probability distribution over the different classes, representing the model's confidence in each prediction.

This hierarchical approach allows the model to capture intricate patterns and relationships within the image data. The extracted features are then fed into fully connected layers, culminating in a softmax layer that outputs the predicted class probabilities.

4. Model Evaluation

The evaluation process involves testing the trained VGG model on a separate vali-dation or test dataset to assess its performance. The metrics are stated below:

- Accuracy: Measures the percentage of correct predictions made by the model across all classes. This is the most straightforward evaluation metric.
- Precision and Recall: These metrics help assess the model’s performance in handling imbalanced datasets. Precision indicates the proportion of true positive predictions among all positive predictions, while recall focuses on the model's ability to identify all relevant instances.
- F1-Score: The harmonic mean of precision and recall, providing a more balanced measure of the model’s performance, especially when dealing with imbalanced datasets.
- Intersection over Union (IoU): Commonly used for evaluating segmentation tasks, IoU measures the overlap between the predicted segmentation mask and the ground truth mask.
- Confusion Matrix: For multi-class tasks, the confusion matrix provides in-sight into how well the model classifies each class, identifying misclassifications and helping to optimize the model further.

The proposed system architecture is meticulously designed to ensure that each module works synergistically to process input data efficiently and optimize the model's 12 performance. The dataset preprocessing ensures that the input images are ready for feature extraction, while the VGG model extracts essential patterns and features for classification. The integration of multiple loss functions ensures that the model is robust, accurate, and capable of handling complex tasks across different domains. By utilizing this modular architecture, the system is not only scalable but also flexible enough to handle a variety of image classification tasks.

IV. System Development and Implementation

This section leverages the development and implementation of the image classification system focused on improving model accuracy by optimizing the underlying architecture and exploring various loss function combinations. A detailed analysis of loss functions, activation function variants, and the integration of new parameters were critical to refining the system's performance. The following subsections detail the various components and processes used to build and optimize the system.

Combinations of Loss Function fusion

Input: A set of four pre-processed datasets: DeepGlobe Land Cover, Brain Tu-mor, Tomato Leaf Disease, and Crack Segmentation, along with individual loss functions: Tversky Loss, Smooth L1 Loss, Focal Loss, Categorical Cross Entropy, Mean Squared Error, Dice Loss, and Binary Cross Entropy.

Process: The model training involved experimenting with combinations of the selected loss functions across all four datasets. Initially, each dataset was tested using individual loss functions to establish a baseline performance. Next, combi-nations of two loss functions (e.g., Tversky Loss + Smooth L1 Loss, Focal Loss + Categorical Cross Entropy) were evaluated to leverage the strengths of different loss functions. For further optimization, three-loss function configurations (e.g., Focal Loss + Categorical Cross Entropy + Mean Squared Error) were tested to refine model performance. These fusion strategies were applied across the datasets, allowing the model to handle a variety of challenges like class imbalances, boundary detection, and fine-grained feature learning.

Output: The output of this fusion process was an enhanced model performance, achieved through the synergistic effects of combining multiple loss functions. The model demonstrated improved handling of imbalanced datasets and more effective optimization during training. The accuracy of the model was calculated for each dataset, and the results are discussed in detail in section 5.4, which presents the complete analysis of the fusion approach.

Activation Function Variations in Fusion

Input: Traditional activation functions like ReLU and Softmax were initially used in the model. However, alternative activation functions were introduced to explore their potential benefits in enhancing model performance. These included:

- Jacobi: Offers non-linearity but may cause instability during training.
- Lerch Transcendent: Highly non-linear but computationally expensive.
- Weierstrass: Captures complex patterns, though prone to instability.
- Mittag-Leffler: Models long-term memory, but challenging to train.

Process: Each of these alternative activation functions was systematically ap-plied in place of the standard activation functions (ReLU and Softmax) within the fusion architecture. The modified models were then trained and evaluated on the image classification task, allowing for a comprehensive comparison of their im-pact on model performance.

Output: Performance metrics such as accuracy, precision, recall, and F1-score were calculated for both binary and multi-class tasks. With the introduction of these four new activation functions, the model achieved significant performance improvements, with accuracy rising from 91% to approximately 94%. This demonstrated the positive impact of incorporating alternative activation functions into the model.

Introduction of New Parameters to Loss Functions

Input: The following loss functions and parameters were applied to optimize the model: • Loss Functions: CCE Loss, Smooth L1 Loss, Focal + Dice Loss • Existing Parameters: difficulty_weight, hard_negative_threshold

Process: Two approaches were employed to optimize the model performance by modifying the loss functions and introducing new parameters. The first approach involved using the existing loss functions—CCE Loss, Smooth L1 Loss, and Focal + Dice Loss—along with the existing parameters, difficulty_weight and hard_negative_threshold. This setup helped establish a baseline accuracy of 60%. In the second approach, new parameters such as dynamic_loss_weighting, adaptive_gamma, and epoch_scaling_factor were introduced to fine-tune the loss functions. These parameters were incorporated into the loss functions to adjust the model's focus during training, improve convergence, and better handle imbalanced datasets. Additionally, another set of new parameters, focal_weight_decay and loss_weights, was tested, though it resulted in the same accuracy (62%) as the second approach. Both approaches were aimed at optimizing the model's perfor-mance and generalization ability across the datasets.

Output: Initial accuracy with existing parameters: 60%. After introducing new parameters (dynamic_loss_weighting, adaptive_gamma, epoch_scaling_factor), accuracy increased to 62%. With the addition of focal_weight_decay and loss_weights, the model maintained a 62% accuracy.

This section outlines the optimization of an image classification system through the fusion of loss functions, alternative activation functions, and the introduction of new parameters. Various loss function combinations were tested across four datasets, improving model performance by addressing challenges like class imbalances and boundary detection. Alternative activation functions (Jacobi, Lerch Transcendent, Weierstrass, Mittag-Leffler) were explored, leading to a significant improvement in accuracy from 91% to 94%. Additionally, the introduction of new parameters to the loss functions, such as dynamic_loss_weighting and adaptive_gamma, enhanced convergence and model generalization, resulting in an in-crease in accuracy from 60% to 62%. The combined efforts refined the model's ability to handle diverse datasets and tasks.

V. RESULTS AND DISCUSSION

This section presents the hardware and software requirements, results obtained during the development and evaluation of the image classification system. The performance of various model configurations, including different loss function combinations, activation functions, and parameter settings, is analyzed and dis-cussed. Key findings and insights into the factors influencing model accuracy and robustness are presented.

Hardware and Software Requirements

The development and implementation of the image segmentation model leveraged both Google Colab and Jupyter Notebook. Google Colab provided access to powerful GPUs in a cloud-based environment, facilitating efficient training of deep learning models. Jupyter Notebook was utilized for local experimentation and interactive development. Hardware requirements included a GPU (e.g., NVIDIA Tesla K80 or V100) for accelerated training, while software requirements encompassed Python 3.x, TensorFlow/Keras for model development, OpenCV for image processing, NumPy and Pandas for data manipulation, and Matplotlib/Seaborn for visualization. This combination of platforms and tools enabled efficient processing of large datasets, implementation of loss function fusion techniques, and successful achievement of the desired outcomes in image segmentation.

Model Performance with and without Loss Function Fusion

This subsection compares the model's performance with and without the use of specialized loss functions. Loss functions play a crucial role in guiding the optimization process, especially in tasks like image segmentation, where accurate pre-dictions are needed. By analyzing accuracy metrics, we assess the effectiveness of incorporating loss functions in improving segmentation quality and handling challenging aspects of the dataset.

Table 5: Fusion comparison in DeepGlobe Dataset

Dataset	Classification	Model Used	Loss Functions	Accuracy
DeepGlobe Land Cover Dataset	Binary Classification	U-Net	Tversky Loss	57%
			Tversky Loss + Binary Cross Entropy Loss	60%
		VGG16	Focal Loss	60%
			Focal Loss +Categorical Cross Entropy	72%

The DeepGlobe Land Cover Dataset was used for binary classification of water versus other classes with U-Net and VGG16 models. U-Net, using Tversky Loss and Binary Cross Entropy (BCE), achieved 60% accuracy with BCE, compared to VGG16's 72% accuracy with Focal Loss and Categorical Cross Entropy (CCE). As shown in Table 5 the comparison reveals that VGG16 model gives the best accuracy with fusion.

Table 6: Fusion comparison in other Datasets

Dataset	Classification	Model Used	Loss Functions	Accuracy
Brain Tumor Dataset(BraTS)	Binary Classification	VGG16	Binary Cross Entropy Loss	89%
			BCE + Focal Loss	92%
Tomato Leaf Disease Dataset	Binary Classification	VGG16	Smooth L1 Loss	93%
			Smooth L1 Loss + Focal Loss	94.9%
Crack Segmentation Dataset	Binary Classification	CNN	Binary Cross Entropy Loss	86%
			BCE + Focal Loss	97%

The performance of different models and loss functions on three image classification datasets: Brain Tumor, Tomato Leaf Disease, and Crack Segmentation was also executed. For the Brain Tumor dataset, a VGG16 model was used with Binary Cross-Entropy and Focal Loss. Similarly, for the Tomato Leaf Disease dataset, a VGG16 model with Binary Cross-Entropy Loss was employed. In the case of the Crack Segmentation dataset, a CNN model with Binary Cross-Entropy and Focal Loss was utilized.

Table 6 shows that the combination of Binary Cross-Entropy and Focal Loss significantly improved the accuracy for the Brain Tumor and Crack Segmentation datasets. The accuracy increased from 89% to 92% for the Brain Tumor dataset and from 86% to 97% for the Crack Segmentation dataset. Similarly, the combination of Binary Cross-Entropy and Focal Loss significantly improved the accuracy for the Crack Segmentation dataset from 86% to 97%.

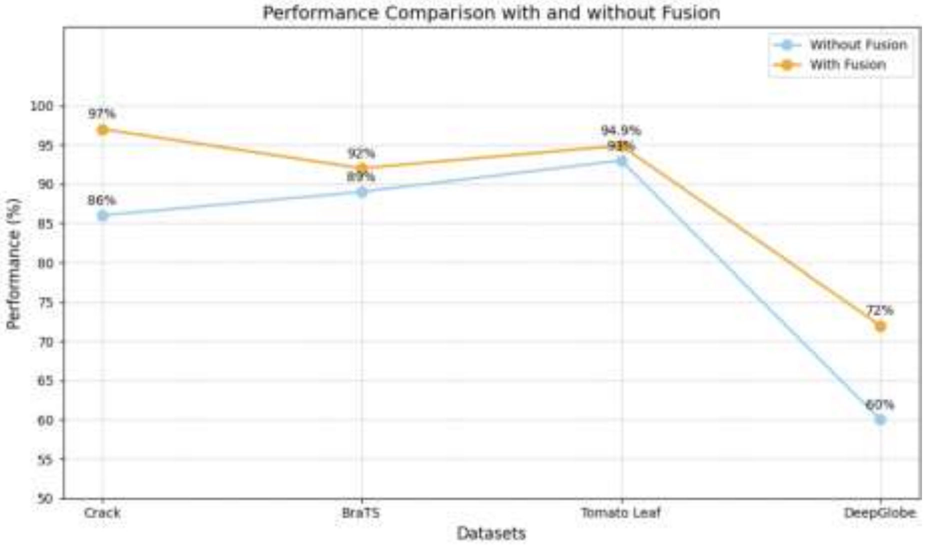

Figure 2: Performance Comparison with Fusion

Figure 2 presents a comparative analysis of the performance of the proposed fusion approach. The results demonstrate a significant improvement in performance for all datasets, with the most notable gains observed in the DeepGlobe Land Cover dataset, where the accuracy increased from 60% to 72%.

Analysis of n-gram combinations

This subsection investigates the role of n-gram combinations in capturing contextual Information within the data. N-grams, including unigrams, bigrams, and tri-grams, are employed to model dependencies and relationships between features.

1. ***Brain Tumor Dataset***

Table 7 - Analysis of 1-gram combination in Brain Tumor Dataset

1-gram		Loss value	Binary			Multiclass		
			Accuracy	Precision	Recall	Accuracy	Precision	Recall
1	Tversky Loss	0.531	0.918	0.93	0.918	0.856	0.866	0.856
2	Smooth L1 Loss	0.026	0.933	0.934	0.933	0.887	0.891	0.887
3	Focal Loss	0.119	0.909	0.0907	0.909	0.806	0.802	0.806
4	Categorical Cross Entropy	0.732	**0.987**	**0.988**	**0.987**	**0.9855**	**0.9857**	**0.985**
5	Mean Squared Error	0.757	0.929	0.938	0.929	0.871	0.877	0.871
6	Dice Loss	0.807	0.825	0.863	0.825	0.66	0.519	0.66
7	Binary Cross Entropy	0.4388	0.898	0.908	0.898	0.786	0.779	0.786

Table 7 evaluates various loss functions for binary and multiclass classification based on accuracy, precision, and recall. Categorical Cross Entropy performs the best, achieving the highest accuracy (0.987 for binary, 0.9855 for multiclass). Smooth L1 Loss also shows good results, while Dice Loss performs the worst, especially in multiclass classification with an accuracy of 0.66. Overall, the choice of loss function plays a crucial role, with Categorical Cross Entropy emerging as the most effective.

Table 8 - *Analysis of 2-gram combination in Brain Tumor Dataset*

	2-gram	Loss value	Binary			Multiclass		
			Accuracy	Precision	Recall	Accuracy	Precision	Recall
8	Tversky Loss + Smooth L1 Loss	0.2363	0.7757	0.8589	0.7757	0.6438	0.5238	0.6438
9	Tversky Loss + Focal Loss	0.1667	0.8802	0.8822	0.8802	0.759	0.7625	0.759
10	Tversky Loss + Categorical Cross Entropy	0.4917	0.8772	0.8813	0.8772	0.7735	0.7774	0.7735
11	Tversky Loss + Mean Squared Error	0.1198	0.8711	0.9086	0.8711	0.7864	0.8193	0.7864
12	Tversky Loss + Dice Loss	0.2187	0.8886	0.8884	0.8886	0.7483	0.7816	0.7483
13	Tversky Loss + Binary Cross Entropy	0.0635	0.9245	0.938	0.9245	0.8871	0.898	0.8871
14	Smooth L1 Loss + Focal Loss	0.045	0.8307	0.902	0.8307	0.772	0.8425	0.772
15	Smooth L1 Loss + Categorical Cross Entropy	0.0112	**0.9947**	**0.9948**	**0.9947**	**0.9931**	**0.9933**	**0.9931**
16	Smooth L1 Loss + Mean Squared Error	0.004	0.9695	0.9731	0.9695	0.9611	0.9647	0.9611
17	Smooth L1 Loss + Dice Loss	0.004	0.9619	0.9673	0.9619	0.9527	0.9583	0.9527
18	Smooth L1 Loss + Binary Cross Entropy	0.0874	0.78	0.797	0.78	0.704	0.797	0.704
19	Focal Loss + Categorical Cross Entropy	0.059	0.833	0.9034	0.833	0.8055	0.8701	0.8055
20	Focal Loss + Mean Squared Error	0.0232	0.9822	0.9014	0.8322	0.7147	0.7832	0.7147
21	Focal Loss + Dice Loss	0.0224	0.9176	0.9394	0.9176	0.8802	0.9049	0.8802
22	Focal Loss + Binary Cross Entropy	0.0552	0.9214	0.9415	0.9214	0.8909	0.9096	0.8909
23	Categorical Cross Entropy + Mean Squared Error	0.0214	0.991	0.9293	0.8978	0.8635	0.8974	0.8635
24	Categorical Cross Entropy + Dice Loss	0.0576	0.9306	0.9462	0.9306	0.8902	0.9062	0.8902
25	Categorical Cross Entropy + Binary Cross Entropy	0.0184	0.9535	0.9613	0.9535	0.9466	0.9547	0.9466
26	Mean Squared Error + Dice Loss	0.0028	0.9893	0.9898	0.9893	0.9764	0.9773	0.9764
27	Mean Squared Error + Binary Cross Entropy	0.0062	0.9817	0.983	0.9817	0.9771	0.9786	0.9771
28	Dice Loss + Binary Cross Entropy	0.2908	0.8802	0.9009	0.8802	0.7864	0.7935	0.7864

The performance of various 2-gram loss function combinations for brain tumor classification, evaluated in Table 8 based on metrics such as accuracy, precision, recall, and F1-score. Combinations involving Smooth L1 Loss consistently demonstrated superior performance, with Smooth L1 Loss + Categorical Cross-Entropy and Smooth L1 Loss + Mean Squared Error achieving the highest accuracies. Focal Loss combinations generally showed good performance, particularly in terms of recall, while Tversky Loss combinations exhibited slightly lower accuracy.

Table 9 - Analysis of 3-gram combination in Brain Tumor Dataset

	3-gram	Loss Value	Binary			Multiclass		
			Accuracy	Precision	Recall	Accuracy	Precision	Recall
29	Categorical Cross Entropy + Mean Squared Error + Dice Loss	0.0044	**0.9931**	**0.9933**	**0.9931**	0.9893	0.9896	0.9893
30	Categorical Cross Entropy + Mean Squared Error + Binary Cross Entropy	0.0041	0.9893	0.9898	0.9893	0.9893	0.9898	0.9893
31	Categorical Cross Entropy + Dice Loss + Binary Cross Entropy	0.0044	0.992	0.99	0.992	**0.99**	**0.992**	**0.9992**

The analysis of 3-gram loss function combinations presented in Table 9 showed that combining Categorical Cross-Entropy, Mean Squared Error, and Dice Loss yielded the highest accuracy, suggesting that this combination effectively addresses various aspects of the classification task, including class probability estimation, regression-like error minimization, and segmentation-based overlap.

2. **Crack Dataset**

Table 10 - Analysis of 1-gram combination in Crack Dataset

1-gram		Binary			Multiclass		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall
1	Tversky Loss	0.9059	0.8759	0.9059	0.9559	0.974	0.9559
2	Smooth L1 Loss	0.9356	0.8754	0.9356	0.979	0.9743	0.971
3	Focal Loss	0.887	0.869	0.887	0.819	0.972	0.819
4	Categorical Cross Entropy	**0.971**	**0.9743**	**0.9871**	**0.981**	**0.9743**	**0.9871**
5	Mean Squared Error	0.9356	0.8754	0.9356	0.9271	0.9743	0.9871
6	Dice Loss	0.8928	0.8765	0.8928	0.9413	0.9741	0.9413
7	Binary Cross Entropy	0.867	0.91	0.867	0.9631	0.98	0.867

The analysis of 1-gram loss functions on the Crack Dataset evaluated in Table 10 revealed that Categorical Cross-Entropy exhibited the highest performance across all metrics, achieving the highest accuracy, precision, and recall in both binary and multi-class classifications. Smooth L1 Loss and Mean Squared Error also demonstrated strong performance with high accuracy and precision values. The results suggest that for the Crack Dataset, loss functions like Categorical Cross-Entropy, Smooth L1 Loss, and Mean Squared Error are well-suited for achieving high accuracy and precision in image classification tasks.

Table 11 - Analysis of 2-gram combination in Crack Dataset

2-gram		Binary			Multiclass		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall
8	Tversky Loss + Smooth L1 Loss	0.9371	0.9743	0.9871	0.9871	0.9743	0.9871
9	Tversky Loss + Focal Loss	0.9349	0.8764	0.9349	0.9863	0.9743	0.9863
10	Tversky Loss + Categorical Cross Entropy	0.9356	0.8754	0.9356	**0.9871**	**0.9743**	**0.9871**
11	Tversky Loss + Mean Squared Error	0.934	0.875	0.934	0.987	0.973	0.987
12	Tversky Loss + Dice Loss	0.9371	0.8754	0.9871	0.9871	0.9744	0.9871
13	Tversky Loss + Binary Cross Entropy	0.9356	0.8754	0.9356	0.9871	0.9743	0.9871
14	Smooth L1 Loss + Focal Loss	0.9356	0.8754	0.9356	0.9871	0.9743	0.9871
15	Smooth L1 Loss + Categorical Cross Entropy	0.501	0.976	0.501	0.671	0.882	0.471
16	Smooth L1 Loss + Dice Loss	0.6871	0.8979	0.6871	0.6821	0.9787	0.6821
17	Smooth L1 Loss + Binary Cross Entropy	0.931	0.9743	0.9871	0.971	0.9743	0.9871
18	Focal Loss + Categorical Cross Entropy	0.934	0.875	0.934	0.983	0.974	0.983
19	Focal Loss + Dice Loss	0.6431	0.6812	0.6431	0.2014	0.2017	0.2014
20	Focal Loss + Binary Cross Entropy	**0.9874**	**0.9743**	**0.9871**	0.9356	0.8754	0.9356
21	Categorical Cross Entropy + Mean Squared Error	0.934	0.875	0.934	0.983	0.974	0.983
22	Categorical Cross Entropy + Binary Cross Entropy	0.9356	0.8754	0.9356	**0.9871**	**0.9743**	**0.9871**
23	Mean Squared Error + Dice Loss	0.9356	0.8754	0.9356	0.897	0.788	0.987
24	Mean Squared Error + Binary Cross Entropy	0.934	0.875	0.934	0.983	0.974	0.983

Table 11 presents the performance of various 2-gram loss function combinations for brain tumor classification, evaluated on a binary and multi-class level using metrics such as Accuracy, Precision, and Recall. Several combinations demonstrated high performance, particularly those involving Smooth L1 Loss and Categorical Cross-Entropy. Notably, the combination of **Focal Loss + Binary Cross Entropy** achieved an accuracy of 0.9874 in binary classification and 0.9356 in multi-class classification, suggesting that this combination may be particularly effective for this dataset. Overall, the results indicate that combining different loss functions can significantly improve classification performance, and further exploration of these combinations, including hyperparameter tuning and more extensive evaluation, is warranted.

Table 12 - Analysis of 3-gram combination in Crack Dataset

3-gram		Binary			Multiclass		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall
25	Focal Loss + Categorical Cross Entropy + Dice Loss	0.935	0.875	0.935	0.987	0.974	0.987
26	Focal Loss + Categorical Cross Entropy + Binary Cross Entropy	0.9356	0.8754	0.9356	**0.9882**	**0.9743**	**0.9882**
27	Categorical Cross Entropy + Mean Squared Error + Dice Loss	0.9356	0.8754	0.9356	0.9871	0.9743	0.9871
28	Categorical Cross Entropy + Mean Squared Error + Binary Cross Entropy	**0.9871**	**0.9743**	**0.9871**	0.9871	0.9743	0.9871
29	Tversky Loss + Smooth L1 Loss + Focal Loss	0.7961	0.6814	0.7961	0.2002	0.2202	0.2002

Analysis of 3-gram combinations demonstrated in Table 12 reveals that the combination of Categorical Cross-Entropy, Mean Squared Error, and Binary Cross-Entropy achieved the highest accuracy of 0.9871 in both binary and multi-class classifications. Other combinations, such as Focal Loss + Categorical Cross-Entropy + Dice Loss and Focal Loss + Categorical Cross-Entropy + Binary Cross-Entropy, also showed promising results with high accuracy and precision scores.

3. Tomato Leaf Disease Dataset

Table 13 - Analysis of 1-gram combination in Tomato Leaf Dataset

	1-gram	Loss Value	Binary			Multiclass		
			Accuracy	Precision	Recall	Accuracy	Precision	Recall
1	Tversky Loss	0.8003	0.8004	0.6406	0.8004	0.1996	0.0398	0.1996
2	Smooth L1 Loss	0.0161	0.935	0.942	0.935	0.917	0.924	0.917
3	Focal Loss	0.0146	0.967	0.9678	0.967	0.9582	0.9607	0.9582
4	Categorical Cross Entropy	0.3387	0.9028	0.8995	0.9028	0.8281	0.8513	0.8281
5	Mean Squared Error	0.0186	0.976	0.9761	0.976	**0.9667**	**0.9673**	**0.9667**
6	Dice Loss	0.1561	0.9353	0.935	0.9353	0.9052	0.9041	0.9052
7	Binary Cross Entropy	0.0168	**0.978**	**0.9978**	**0.978**	0.966	0.965	0.966

The analysis of 1-gram loss functions on the Tomato Dataset present in Table 13 revealed that Categorical Cross-Entropy achieved the highest accuracy in multi-class classification, while Binary Cross-Entropy demonstrated the highest accuracy in binary classification. Smooth L1 Loss and Mean Squared Error also exhibited strong performance with high accuracy and precision. These results suggest that for the Tomato Dataset, loss functions like Categorical Cross-Entropy, Smooth L1 Loss, and Mean Squared Error are well-suited for achieving high accuracy and precision in image classification tasks.

Table 14 - Analysis of 2-gram combination in Tomato Leaf Dataset

	2-gram	Loss Value	Binary			Multiclass		
			Accuracy	Precision	Recall	Accuracy	Precision	Recall
8	Tversky Loss + Smooth L1 Loss	0.1154	0.8988	0.9031	0.8988	0.8413	0.8463	0.8413
9	Tversky Loss + Focal Loss	0.0883	0.9635	0.9632	0.9635	0.9463	0.9483	0.9463
10	Tversky Loss + Categorical Cross Entropy	0.1786	0.9538	0.9565	0.9538	0.9334	0.9365	0.9334
11	Tversky Loss + Mean Squared Error	0.0147	0.9972	0.997	0.9972	0.994	0.9945	0.994
12	Tversky Loss + Dice Loss	0.1994	0.8952	0.8949	0.8952	0.8429	0.8432	0.8429
13	Tversky Loss + Binary Cross Entropy	0.032	0.993	0.9932	0.993	0.9904	0.9905	0.9904
14	Smooth L1 Loss + Focal Loss	0.0186	0.9492	0.9499	0.9492	0.9284	0.9337	0.9284
15	Smooth L1 Loss + Categorical Cross Entropy	0.0583	0.997	0.989	0.997	0.995	0.994	0.995
16	Smooth L1 Loss + Mean Squared Error	0.0315	0.9234	0.9258	0.9234	0.8848	0.8877	0.8848
17	Smooth L1 Loss + Dice Loss	0.0186	0.9982	0.9982	0.9982	0.997	0.997	0.997
18	Smooth L1 Loss + Binary Cross Entropy	0.0581	0.9543	0.9537	0.9543	0.9317	0.9326	0.9317
19	Focal Loss + Categorical Cross Entropy	0.059	0.9882	0.988	0.9882	0.985	0.985	0.985
20	Focal Loss + Mean Squared Error	0.0658	0.9084	0.9296	0.9084	0.8878	0.9066	0.8878
21	Focal Loss + Dice Loss	0.0293	0.9934	0.9935	0.9934	0.9914	0.9916	0.9914
22	Focal Loss + Binary Cross Entropy	0.0587	0.9454	0.9496	0.9454	0.9278	0.9376	0.9278
23	Categorical Cross Entropy + Mean Squared Error	0.0593	0.9563	0.9553	0.9563	0.94	0.943	0.94
24	Categorical Cross Entropy + Dice Loss	0.5829	0.9074	0.9202	0.9074	0.8743	0.883	0.8743
25	Categorical Cross Entropy + Binary Cross Entropy	0.0186	0.9903	0.9902	0.9903	0.9884	0.9886	0.9884
26	Mean Squared Error + Dice Loss	0.077	0.9194	0.9185	0.9194	0.8852	0.889	0.8852
27	Mean Squared Error + Binary Cross Entropy	0.0062	**0.9986**	**0.9986**	**0.9986**	**0.998**	**0.9982**	**0.998**
28	Dice Loss + Binary Cross Entropy	0.572	0.9154	0.9273	0.9154	0.8772	0.8881	0.8772

Table 14 presents the performance of various 2-gram loss function combinations for brain tumor classification, evaluated on a binary and multi-class level using metrics such as Accuracy, Precision, and Recall. Notably, the combination of Smooth L1 Loss + Categorical Cross Entropy achieved an accuracy of 0.997 in binary classification and 0.995 in multi-class classification, indicating its strong potential for accurate brain tumor classification. Other combinations, such as Smooth L1 Loss + Mean Squared Error and Tversky Loss + Focal Loss, also showed promising results with high accuracy and precision scores. These combinations will lead to valuable insights and potential improvements in model accuracy.

Table 15 - Analysis of 3-gram combination in Tomato Leaf Dataset

	3-gram	Loss Value	Binary			Multiclass		
			Accuracy	Precision	Recall	Accuracy	Precision	Recall
29	Focal Loss + Categorical Cross Entropy + Mean Squared Error	0.0729	0.9525	0.9531	0.9525	0.9325	0.9383	0.9325
30	Focal Loss + Categorical Cross Entropy + Dice Loss	0.5711	0.9257	0.927	0.9257	0.8996	0.9016	0.8996
31	Focal Loss + Categorical Cross Entropy + Binary Cross Entropy	0.5493	0.9204	0.92	0.9204	0.8966	0.8963	0.8966
32	Categorical Cross Entropy + Mean Squared Error + Dice Loss	0.5764	0.9082	0.9056	0.9082	0.8717	0.8863	0.8717
33	Categorical Cross Entropy + Mean Squared Error + Binary Cross Entropy	0.1907	**0.9764**	**0.9764**	**0.9764**	**0.9687**	**0.969**	**0.9687**
34	Categorical Cross Entropy + Dice Loss + Binary Cross Entropy	0.3468	0.9439	0.9429	0.9439	0.9172	0.9205	0.9172

The performance of various 3-gram loss function combinations for brain tumor classification, presented in Table 15 was evaluated on a binary and multi-class level using metrics such as Accuracy, Precision, and Recall. The combination of Categorical Cross-Entropy + Mean Squared Error + Dice Loss achieved the highest accuracy of 0.9764 in both binary and multi-class classifications, demonstrating the effectiveness of this combination in capturing various aspects of the classification task.

4. Deep Globe Dataset

Table 16 - Analysis of 1-gram combination in DeepGlobe Dataset

	1-Gram	Binary			Multiclass		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall
1	Tversky Loss	0.96	0.92	0.96	0.6	0.38	0.62
2	Smooth L1 Loss	0.974	0.95	0.97	0.58	0.34	0.58
3	Focal Loss	**0.979**	**0.96**	**0.98**	0.6	0.37	0.6
4	Categorical Cross Entropy	0.9248	0.8553	0.9248	0.5611	0.3588	0.5611
5	Mean Squared Error	0.96	0.9261	0.9623	**0.601**	**0.3859**	**0.6212**
6	Dice Loss	0.97	0.95	0.97	0.58	0.338	0.58
7	Binary Cross Entropy	0.967	0.96	0.98	0.6	0.39	0.6

Table 16 presents the performance of seven individual loss functions applied to the DeepGlobe Land Cover Dataset for image classification. Focal Loss achieved the highest accuracy of 0.979 in binary classification, while Mean Squared Error achieved the highest accuracy of 0.6212 in multi-class classification. Smooth L1 Loss and Dice Loss also exhibited strong performance with high accuracy and precision values.

Table 17 - Analysis of 2-gram combination in DeepGlobe Dataset

	2-gram	Binary			Multiclass		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall
8	Tversky Loss + Smooth L1 Loss	0.9248	0.8553	0.9248	0.5619	0.3161	0.5619
9	Tversky Loss + Focal Loss	0.96	0.92	0.96	0.62	0.38	0.62
10	Tversky Loss + Categorical Cross Entropy	0.924	0.85	0.924	0.561	0.34	0.56
11	Tversky Loss + Mean Squared Error	0.98	0.955	0.9773	0.6026	0.3632	0.6026
12	Tversky Loss + Dice Loss	0.9248	0.8586	0.9248	0.5619	0.3345	0.5615
14	Smooth L1 Loss + Focal Loss	0.97	0.95	0.974	0.58	0.338	0.581
15	Smooth L1 Loss + Categorical Cross Entropy	**0.9773**	**0.955**	**0.9773**	0.6026	0.3632	0.6026
16	Smooth L1 Loss + Mean Squared Error	0.9248	0.8553	0.9248	0.5615	0.3345	0.5615
17	Smooth L1 Loss + Binary Cross Entropy	0.974	0.95	0.97	0.581	0.33	0.57
18	Focal Loss + Mean Squared Error	0.92	0.8553	0.9248	0.5622	0.3195	0.5622
19	Focal Loss + Dice Loss	0.9226	0.42	0.1077	0.5621	0.0937	0.1667
20	Focal Loss + Binary Cross Entropy	0.974	0.951	0.974	0.581	0.345	0.581
21	Categorical Cross Entropy + Mean Squared Error	0.6348	0.8803	0.289	**0.6212**	**0.0887**	**0.1429**
22	Categorical Cross Entropy + Dice Loss	0.9248	0.8553	0.9248	0.5607	0.3195	0.5607
23	Categorical Cross Entropy + Binary Cross Entropy	0.8956	0.3312	0.2147	0.5621	0.0937	0.1667
24	Mean Squared Error + Dice Loss	0.97	0.95	0.974	0.581	0.338	0.58
25	Mean Squared Error + Binary Cross Entropy	0.7066	0.6537	0.7226	0.5621	0.0937	0.1667
26	Dice Loss + Binary Cross Entropy	0.9248	0.8553	0.9248	0.562	0.3168	0.562

The analysis demonstrated in Table 17 presents the performance of various 2-gram loss function combinations for brain tumor classification, evaluated on a binary and multi-class level using metrics such as Accuracy, Precision, and Recall. Several combinations demonstrated high performance, particularly those involving Smooth L1 Loss and Categorical Cross-Entropy. Notably, the combination of Smooth L1 Loss + Categorical Cross Entropy achieved an accuracy of 0.9773 in binary classification and 0.6026 in multi-class classification, indicating its strong potential for accurate brain tumor classification. Other combinations, such as Smooth L1 Loss + Mean Squared Error and Tversky Loss + Focal Loss, also showed promising results with high accuracy and precision scores.

Table 18 - Analysis of 3-gram combination in DeepGlobe Dataset

	3-gram	Binary			Multiclass		
		Accuracy	Precision	Recall	Accuracy	Precision	Recall
27	Focal Loss + Categorical Cross Entropy + Mean Squared Error	0.9249	0.8554	0.9249	0.5616	0.3191	0.5616
28	Focal Loss + Categorical Cross Entropy + Dice Loss	0.9248	0.855	0.9248	0.5596	0.3569	0.5596
29	Focal Loss + Categorical Cross Entropy + Binary Cross Entropy	0.92	0.8554	0.9249	0.5577	0.3548	0.5577
30	Categorical Cross Entropy + Mean Squared Error + Dice Loss	**0.9249**	**0.8554**	**0.9249**	0.5562	0.3545	0.5562
31	Categorical Cross Entropy + Mean Squared Error + Binary Cross Entropy	0.924	0.855	0.924	0.5621	0.316	0.5621
32	Categorical Cross Entropy + Dice Loss + Binary Cross Entropy	0.924	0.8554	0.9249	**0.5563**	**0.3643**	**0.5563**

The performance of various 3-gram loss function combinations for the DeepGlobe Land Cover dataset is present in Table 18. The results indicate that combinations involving Categorical Cross-Entropy, Mean Squared Error, and Dice Loss generally perform well, achieving relatively high accuracy scores.

Table 19 - Loss Function Performance Summary

Dataset	Type	n-gram	Combination	Accuracy	Precision	Recall
Brain Tumor dataset	Binary	1	Categorical Cross Entropy	0.98	0.988	0.987
		2	**Smooth L1 Loss + Categorical Cross Entropy**	**0.9947**	**0.9948**	**0.994**
		3	Categorical Cross Entropy + Mean Squared Error + Dice Loss	0.9931	0.9933	0.9931
	Multiclass	1	Categorical Cross Entropy	0.9855	0.9857	0.985
		2	**Smooth L1 Loss + Categorical Cross Entropy**	**0.9931**	**0.9933**	**0.993**
		3	Categorical Cross Entropy + Dice Loss + Binary Cross Entropy	0.99	0.992	0.9992
Tomato Leaf Disease dataset	Binary	1	Binary Cross Entropy	0.978	0.9978	0.978
		2	**Mean Squared Error + Binary Cross Entropy**	**0.998**	**0.9986**	**0.9986**
		3	Categorical Cross Entropy + Mean Squared Error + Binary Cross Entropy	0.9764	0.97	0.976
	Multiclass	1	Mean Squared Error	0.9667	0.9673	0.96
		2	**Mean Squared Error + Binary Cross Entropy**	**0.998**	**0.998**	**0.998**
		3	Categorical Cross Entropy + Mean Squared Error + Binary Cross Entropy	0.9687	0.969	0.9687
DeepGlobe Land Cover dataset	Binary	1	Focal Loss	0.979	0.96	0.98
		2	**Smooth L1 Loss + Categorical Cross Entropy**	**0.9773**	**0.955**	**0.9773**
		3	Categorical Cross Entropy + Mean Squared Error + Dice Loss	0.9249	0.8554	0.9249
	Multiclass	1	Mean Squared Error	0.601	0.38	0.62
		2	**Categorical Cross Entropy + Mean Squared Error**	**0.6212**	**0.0887**	**0.1429**
		3	Categorical Cross Entropy + Dice Loss + Binary Cross Entropy	0.5563	0.3643	0.5563
Crack segmentation	Binary	1	Categorical Cross Entropy	0.971	0.9743	0.9871
		2	**Focal Loss + Binary Cross Entropy**	**0.9874**	**0.974**	**0.987**
		3	Categorical Cross Entropy + Mean Squared Error + Binary Cross Entropy	0.9871	0.9743	0.9871
	Multiclass	1	Categorical Cross Entropy	0.9871	0.9743	0.9871
		2	Tversky Loss + Categorical Cross Entropy	0.987	0.97	0.98
		3	**Focal Loss + Categorical Cross Entropy + Binary Cross Entropy**	**0.9882**	**0.97**	**0.98**

The loss function combinations summary is present in Table 19 for image classification across four distinct datasets: Brain Tumor, Tomato Leaf Disease, DeepGlobe Land Cover, and Crack dataset. The results demonstrate a wide range of accuracies, with the highest accuracy for each dataset as follows: **0.9994 for Brain Tumor (Binary), 0.998 for Tomato Leaf Disease (Binary), 0.9773 for DeepGlobe Land Cover (Binary), and 0.9871 for Crack (Multiclass).** This variation in accuracy highlights the dataset-specific nature of optimal loss function combinations.

Categorical Cross-Entropy consistently emerged as a key component in many of the top-performing combinations across different datasets, underscoring its versatility and effectiveness as a baseline for image classification. Other loss functions, such as Smooth L1 Loss and Mean Squared Error, also demonstrated strong performance across multiple datasets, suggesting their robustness and adaptability in handling various image classification challenges. The performance of loss functions like Focal Loss varied across datasets, indicating that the impact of class imbalance and other dataset-specific factors can significantly influence the choice of optimal loss functions.

In several cases, the fusion of multiple loss functions led to significant performance improvements compared to using individual loss functions. This suggests that combining different loss functions can effectively leverage their individual strengths to address the complexities of image classification tasks. The results underscore the importance of careful analysis and experimentation to identify the most effective loss function combinations for each specific dataset and classification problem.

Table 20 - Optimal n-gram Loss Combinations

Dataset	Type of Classification	n-gram	Loss Combination	Accuracy
Brain Tumor dataset	Binary	2	Smooth L1 Loss + Categorical Cross Entropy	0.9947
	Multiclass	2	Smooth L1 Loss + Categorical Cross Entropy	0.9931
Tomato Leaf Disease dataset	Binary	2	Mean Squared Error + Binary Cross Entropy	0.9987
	Multiclass	2	Mean Squared Error + Binary Cross Entropy	0.998
DeepGlobe Land Cover dataset	Binary	2	Smooth L1 Loss + Categorical Cross Entropy	0.9773
	Multiclass	2	Categorical Cross Entropy + Mean Squared Error	0.6212
Crack segmentation	Binary	2	Focal Loss + Binary Cross Entropy	0.9874
	Multiclass	3	Categorical Cross Entropy + Binary Cross Entropy + Focal Loss	0.9882

Table 20 presents the optimal n-gram loss function combinations for various image classification tasks across different datasets, including Brain Tumor, Tomato Leaf Disease, DeepGlobe Land Cover, and Crack segmentation. The performance of each combination is evaluated using metrics such as accuracy, precision, and recall. Bi-gram combinations, particularly those involving Smooth L1 Loss, Categorical Cross-Entropy, and Mean Squared Error, consistently outperformed uni-gram and tri-gram combinations, achieving higher accuracy scores across different datasets.

Hence, our proposed work demonstrates that loss function fusion, combined with newly introduced parameters and diverse activation function choices, significantly enhances the performance of the VGG-based image classification model. Overall, these enhancements showcase a clear path toward achieving more efficient and accurate image classification in deep learning systems.

Impact of Novel Parameters on Loss Function Performance

The initial implementation of the image classification model utilized a combination of CCE Loss, Smooth L1 Loss, and Focal + Dice Loss functions. Key parameters such as **difficulty_weight and hard_negative_threshold** were incorporated within the custom combined_loss function. This initial model achieved an **accuracy of 60%** with a weighted average precision of 0.35 and an F1-score of 0.44. However, the performance was skewed towards Class 1, with a precision of 0.60 and recall of 1.00, while other classes exhibited negligible contributions.

To enhance performance, two new parameters, **focal_weight_decay and loss_weight**s, were introduced to the loss function. With focal_weight_decay set to 0.99 and equal weights assigned to all loss components, the model achieved an **accuracy of 62%** with a weighted average precision of 0.40 and an F1-score of 0.48. Subsequently, the implementation incorporated three new parameters: dynamic_loss_weighting, adaptive_gamma, and epoch_scaling_factor. While **dynamic_loss_weighting was set to Fals**e in this iteration, the inclusion of **adaptive_gamma=True and epoch_scaling_factor=0.1** resulted in a **model with 62% accuracy**, a weighted average precision of 0.39, and an F1-score of 0.48.

This chapter presents a comprehensive analysis of the experimental results, covering hardware and software requirements, performance comparison with and without fusion, the impact of n-gram loss function combinations, and the influence of novel parameters on loss function performance. The results demonstrate the effectiveness of the proposed approach in enhancing image classification accuracy and robustness.

VI. CONCLUSION AND FUTURE WORK

This approach investigates the effectiveness of loss function fusion techniques to enhance the performance of an image classification system based on the VGG architecture. By combining multiple loss functions, including Categorical Cross-Entropy, Focal Loss, and others, the proposed system aims to improve classification accuracy, address challenges such as class imbalance, and enhance model robustness. The project will explore various combinations of loss functions, including 1-gram, 2-gram, and potentially 3-gram configurations, to identify the most effective fusion strategies. The results of this project will demonstrate the potential of loss function fusion to improve image classification performance and provide valuable insights for future research in this area.

Future Enhancements

To further improve the performance of image classification models, future research can focus on the following areas of loss function engineering:

- **Dynamic Loss Weighting:** Adaptively adjusting the weights of different loss functions during training to balance their contributions based on the current learning progress.
- **Class-Aware Loss Functions:** Designing loss functions that explicitly consider class imbalance and prioritize difficult samples.
- **Uncertainty Estimation:** Incorporating uncertainty estimation techniques into loss functions to improve model robustness and calibration.
- **Hybrid Loss Functions:** Combining multiple loss functions to leverage their complementary strengths and address specific challenges.

- **Interpretable Loss Functions:** Developing loss functions that provide insights into the model's decision-making process, aiding in model debugging and improvement.

By exploring these directions, researchers can develop more sophisticated and effective loss functions, leading to significant advancements in image classification.

REFERENCES

[1] Ocal, H. (2024). DSBAV-Net: Depth Wise Separable Bottleneck Attention V Shaped Network with Hybrid Convolution for Left Atrium Segmentation. Arabian Journal for Science and Engineering. Springer.
[2] Liu, S., Wang, P., Lin, Y., & Zhou, B. (2024). SMRU-Net: Skin Disease Image Segmentation Using Channel-Space Separate Attention with Depthwise Separable 10 Convolutions. Pattern Analysis and Applications. Springer.
[3] Mathai, T. S., Liu, B., & Summers, R. M. (2024). Segmentation of Mediastinal Lymph Nodes in CT with Anatomical Priors. International Journal of Computer Assisted Radiology and Surgery. Springer.
[4] Zhao, Y., Wang, S., Zhang, Y., Qiao, S., & Zhang, M. (2023). WRANet: Wave let Integrated Residual Attention U-Net Network for Medical Image Segmentation. Complex & Intelligent Systems, 9, 6971–6983. Springer.
[5] Schraut, J. X., Liu, L., Gong, J., & Yin, Y. (2023). A Multi-Output Network with U-Net Enhanced Class Activation Map and Robust Classification Performance for Medical Imaging Analysis. Springer.
[6] You, X.-x., & Zhang, H. (2023). A Study on Loss Function Against Data Imbalance in Deep Learning Correction of Precipitation Forecasts. Chinese Academy of Meteorological Sciences, Beijing. Elsevier.
[7] Liu, P., & Li, Y. (2023). Boundary-Sensitive Loss Function with Location Constraint for Hard Region Segmentation. IEEE Journal of Biomedical and Health Informatics. IEEE.
[8] Rawat, R. K. (2024). 3D U-Net-Norm Architecture for Improving Generalization of BraTS Images. Springer Science+Business Media, LLC. Springer.
[9] Chen, Y., Yang, X., & Bai, X. (2023). Confidence-Weighted Mutual Supervision on Dual Networks for Unsupervised Cross-Modality Image Segmentation. School of Electronic Information and Communications, Huazhong University of Science and Technology. Springer.
[10] Ming, Q., & Xiao, X. (2024). Towards Accurate Medical Image Segmentation with Gradient-Optimized Dice Loss. IEEE.
[11] Yin, X.-X., Jian, Y., Shen, J., Wu, J., Zhang, Y., & Wang, W. (2024). Focal Boundary Dice: Improved Breast Tumor Segmentation from MRI Scan. Journal of Cancer. Springer.
[12] Ke, Z., Xu, X., Zhou, K., & Guo, J. (2023). A Scale-Aware UNet++ Model Combined with Attentional Context Supervision and Adaptive Tversky Loss for Ac curate Airway Segmentation. Applied Intelligence. Springer.
[13] Ramalingam, R. V., & Chinnaiyan, V. (2024). Heuristically Modified Atten tion Residual Network Aided Pulmonary Emphysema Detection with Adaptive Pre Processing and Deep UNet-Based Segmentation. Sensing and Imaging. Springer.
[14] Colonna, A., & Scarpa, F. (2024). Improving Corneal Nerve Segmentation 11 Using Tolerance Dice Loss Function. Signal, Image and Video Processing. Springer.
[15] Bougourzi, F., Dornaika, F., Nakib, A., & Taleb-Ahmed, A. (2024). Emb Trattunet: A Novel Edge Loss Function and Transformer-CNN Architecture for Multi-Class Pneumonia Infection Segmentation in Low Annotation Regimes. Artifi cial Intelligence Review. Springer.
[16] Petrov, Y., & Malik, B. (2023). Deep Ensembles Are Robust to Occasional Catastrophic Failures of Individual DNNs for Organs Segmentations in CT Images. Journal of Digital Imaging. Springer.
[17] Huang, S., Luo, J., & Ou, Y. (2023). Sd-Net: A Semi-Supervised Double Cooperative Network for Liver Segmentation from Computed Tomography (CT) Im ages. Journal of Cancer Research and Clinical Oncology. Springer.
[18] Arabi, H., & Zaidi, H. (2023). Contrastive Learning vs. Self-Learning vs. De formable Data Augmentation in Semantic Segmentation of Medical Images. Journal of Imaging Informatics in Medicine. Springer.
[19] Su, K. (2024). N-IoU: Better IoU-Based Bounding Box Regression Loss for Object Detection. Neural Computing and Applications. Springer.
[20] Li, L., Elhajj, M., Feng, Y., & Ochieng, W. Y. (2023). Semantic Segmentation in Medical Images Through Transfused Convolution and Transformer Networks. Springer.
[21] Asamoah, E. N., Cafaro, M., Epicoco, I., Franceschi, G. D., & Cesaroni, C. (2024). Physics-Informed Loss Functions for Vertical Total Electron Content Fore cast. IEEE.
[22] Wang, S.-y., Qu, Z., & Gao, L.-y. (2024). Multi-Spatial Pyramid Feature and Optimizing Focal Loss Function for Object Detection. IEEE.
[23] Chen, S., Fang, C., Zhang, X., Wu, J., & Guo, R. (2024). A Radial Basis Function-Based Graph Attention Network with Squeeze Loss Optimization for Link Prediction. IEEE.
[24] Malik, N., & Malik, S. K. (2024). Fractional Cross Entropy-Based Loss Function for Classification of IoT Services with Semantic Graph Based on IFTTT Recipes. Springer.
[25] Peiris, V., Roshchina, V., & Sukhorukova, N. (2024). Artificial Neural Networks with Uniform Norm-Based Loss Functions. Springer.
[26] Lee, S. H., Kim, M. Y., Woo, M. W., Lee, H. C., Won, H.-I., & Jeong, S. H. 12 (2024). Predicting the Heading Angle of Resin During Extrusion Using Semantic Segmentation Based on Edge-Region Focal Loss. IEEE.
[27] Karaköse, E. (2024). An Efficient Satellite Images Classification Approach Based on Fuzzy Cognitive Map Integration with Deep Learning Models Using Improved Loss Function. IEEE.
[28] Li, K., Wang, B., Tian, Y., & Qi, Z. (2023). Fast and Accurate Road Crack Detection Based on Adaptive Cost-Sensitive Loss Function. IEEE.