B M S COLLEGE OF ENGINEERING

(An Autonomous Institution Affiliated to VTU, Belagavi)

Post Box No.: 1908, Bull Temple Road, Bengaluru – 560 019

DEPARTMENT OF MACHINE LEARNING

Academic Year: 2023-2024 (Session: April 2024 - August 2024))

COURSE: DEEP LEARNING (24AM5PCDEL)

Alternative Assessment Tool (AAT)

Submitted by

Student Name:	**M S Anagha** **Shriya Sai**
Branch:	**Artificial intelligence and Machine learning**
USN:	**1BM22AI067** **1BM22AI126**
Date:	**13-01-2025**
Semester & Section:	**5B**
Total Pages:	
Student Signature:	

Valuation Report **(to be filled by the faculty)**

Score:	
Comments:	
Faculty In-charge:	**Dr.Gowrishankar**
Faculty Signature: with date	

Questions to be addressed

1] Transfer Learning with ResNet-50

Q:

Why is ResNet-50 a good choice for this task? If the model underperforms, what criteria would you use to select an alternative backbone? Provide two suitable alternatives and justify your choices.

A:

ResNet-50 is widely used for transfer learning because:

1. **Residual Connections**: These connections mitigate the vanishing gradient problem, enabling the network to train effectively even at significant depths.
2. **Pre-trained Availability**: ResNet-50 has pre-trained weights on large datasets like ImageNet, making it highly suitable for transfer learning across a variety of tasks.
3. **Balanced Complexity**: With 50 layers, it offers a good trade-off between model capacity and computational cost.

Criteria for Selecting an Alternative Backbone:

1. **Task-Specific Performance**: How well the backbone architecture extracts features relevant to the target task (e.g., fine-grained features for classification or object detection).
2. **Computational Resources**: Model efficiency in terms of memory and processing power requirements.
3. **Dataset Size**: Larger datasets can handle more complex models, while smaller datasets benefit from simpler architectures to prevent overfitting.

Two Suitable Alternatives and Justifications:

1. **DenseNet-121**:

   DenseNet connects each layer to every other layer, promoting feature reuse, which can be advantageous for tasks requiring fine-grained feature extraction.

2. **EfficientNet-B3**:
   - **Accuracy vs. Efficiency Trade-off**: EfficientNet scales depth, width, and resolution systematically, resulting in better performance with fewer parameters compared to traditional architectures.
   - **Scalability**: It offers a range of versions (B0 to B7), making it adaptable to available resources and dataset characteristics.

These alternatives can address limitations in feature representation or computational constraints observed with ResNet-50.

2] Handling Data Imbalance

Q:

The dataset has significantly fewer samples for rare diseases like tuberculosis. Despite weighted loss functions, the model shows low sensitivity for rare diseases. Propose additional strategies to address this issue and evaluate their trade-offs.

A:

Strategies for Handling Data Imbalance:

1. Data Augmentation:

Apply transformations (e.g., rotation, scaling, flipping) to increase the variety of samples for rare classes.

2. Over-Sampling:

Use techniques like Synthetic Minority Oversampling Technique (SMOTE) or its variants to create synthetic samples for underrepresented classes.

3. Active Learning:

Use model predictions to identify misclassified or uncertain samples, then label these cases with expert input.

4. Cost-Sensitive Learning:

Adjust loss functions to penalize misclassifications of rare classes more heavily (e.g., focal loss).

5. Transfer Learning with Domain-Specific Data:

Fine-tune the model using additional external datasets specific to rare diseases.

6. Ensemble Methods:

Combine predictions from multiple models trained on different subsets or re-weighted datasets.

By combining these strategies, particularly data augmentation or over-sampling with cost-sensitive learning, sensitivity for rare diseases like tuberculosis can improve while mitigating overfitting or computational burdens.

3] Heatmap Interpretability

Q:

Grad-CAM is used for generating heatmaps. Explain how Grad-CAM enhances trust in the model’s predictions and discuss its limitations. Suggest another method for interpretability and compare it with Grad-CAM.

A:
Grad-CAM (Gradient-weighted Class Activation Mapping) generates class-specific heatmaps by leveraging the gradients of the target class with respect to the feature maps in a convolutional layer. It:

1. **Highlights Key Regions**: Identifies image regions most influential for the model's prediction, helping users understand the basis of the decision.

2. **Improves Transparency**: Shows if the model focuses on relevant areas, boosting confidence in its correctness.

Limitations of Grad-CAM:

1. **Noisy Heatmaps**: The visualizations can be imprecise, especially for complex backgrounds or overlapping objects.

2. **Poor Localization**: It often struggles to pinpoint small, fine-grained regions (e.g., in medical imaging for subtle anomalies).

3. **Dependence on Convolutional Layers**: Grad-CAM relies on convolutional layers, making it less suited for models without these components.

Alternative Method: LIME (Local Interpretable Model-Agnostic Explanations)

LIME explains individual predictions by approximating the model's behaviour locally. It perturbs the input (e.g., masking parts of an image) and evaluates the model's response to identify influential regions.

4] Evaluation Metrics

Q:

Abnormality detection and severity classification have different objectives. What metrics would you use for each task, and why? How would you address conflicts between optimizing these metrics during training?

A:

1. Abnormality Detection (Binary Classification):

- **Metrics**: Precision, Recall, F1-score (balance false positives/negatives), ROC-AUC (threshold-independent).

2. Severity Classification (Regression/Classification):

- **Metrics**:
  - Regression: MSE (penalizes large errors), MAE (robust to outliers).
  - Multi-class: Cohen's Kappa (accounts for chance agreement).

Conflict Resolution:

- Use **multi-objective loss**: Combine detection and classification losses with weights

$$\text{Loss}=\alpha\times\text{Detection Loss}+\beta\times\text{Classification Loss}$$

  Adjust α and β to prioritize tasks based on their importance.

- **Task-specific heads**: Shared backbone with separate detection/classification outputs.
- **Separate models**: When tasks conflict significantly.

Start with multi-task learning, adjusting priorities with weighted losses. Shift to separate models if conflicts persist.

5] Real-World Constraints

Q:

In real-world deployment, the model needs to process a large volume of X-rays daily under strict latency requirements. How would you optimize the system to meet these constraints without sacrificing accuracy?

A:

1. Model Compression:

- Techniques:
  - Pruning: Remove redundant weights or neurons.
  - Quantization: Use lower precision (e.g., INT8) instead of full precision (FP32).
  - Knowledge Distillation: Train a smaller model (student) to mimic a larger model (teacher).
- Impact: Reduces model size and inference time with minimal accuracy loss if done carefully.

2. Edge Computing:

- Deploy the model on local devices (e.g., edge servers) near data sources.
- Impact: Minimizes latency by avoiding network delays associated with cloud processing.

3. Batch Processing:

- Process multiple images in parallel during inference.
- Leverage optimized hardware like GPUs or TPUs for high throughput.
- Impact: Increases overall efficiency but requires managing memory constraints carefully.

4. Hardware Optimization:

- Use hardware accelerators (e.g., GPUs, TPUs, FPGAs) optimized for deep learning workloads.
- Ensure software frameworks (e.g., TensorRT) are tuned for the deployment hardware.

5. Pipeline Optimization:

- Pre-process images efficiently (e.g., resizing, normalization).
- Use asynchronous processing to overlap I/O and computation tasks.

Recommendation:
Combine model compression with edge computing for low-latency deployment. Use batch processing and hardware optimization to handle high throughput while maintaining accuracy.