	Pimpri Chinchwad Education Trust's **Pimpri Chinchwad College of Engineering (PCCoE)** (An Autonomous Institute) Affiliated to Savitribai Phule Pune University(SPPU) ISO 21001:2018 Certified by TUV SUD	
	Formative Assessment – 2 Report	

FORMATIVE ASSESSMENT -2 REPORT

Institute Name: Pimpri Chinchwad College of Engineering (PCCoE)

Department: Information Technology

Subject: Deep Learning

Academic Year: 2025–26

Title: Formative Assessment Report – Deep Learning Mini Project

Student Name: Harivdan Shrihari Narayanshastri

PRN: 125M1K010

Faculty Name: Prof. Tanuja Patankar

Course / Class: F.Y. M.Tech (AI & DS)

Submission Date:15 DEC 2025

Contents

1. Formative Assessment 2

1.1 Objective

The objective of this assessment is to implement and evaluate two deep learning models proposed in FA1 for COVID-19 medical image classification and demonstrate experimental innovation with measurable improvements. The specific objectives are:

- Implement Model A (CNN Baseline) using EfficientNetB3 with transfer learning.
- Implement Model B (Hybrid CNN-Transformer) incorporating self-attention and transformer encoder layers.
- Compare performance using accuracy, precision, recall, F1-score, and AUC-ROC.
- Achieve at least 90% classification accuracy across four disease categories.
- Validate FA1 hypothesis that hybrid CNN-Transformer models outperform pure CNNs.

1.2 Model Implementation

1.2.1 Dataset Selection and Preparation

Dataset: COVID-19 Radiography Database
Source: `https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database`

Dataset Composition

Preprocessing Steps

1. Image resizing to 224×224 pixels.
2. RGB color conversion and normalization to $[0, 1]$.
3. Stratified dataset split: 70% training, 15% validation, 15% testing.
4. Class imbalance handled using weighted categorical cross-entropy.

Class	Number of Images
COVID-19	3,616
Lung Opacity	6,012
Normal	10,192
Viral Pneumonia	1,345
Total	21,165

Table 1.1: Dataset Distribution

Framework and Tools

- TensorFlow 2.15.0 with Keras
- Python 3.10
- Google Colab with NVIDIA Tesla T4 GPU
- Libraries: NumPy, Pandas, OpenCV, Scikit-learn, Matplotlib

1.2.2 Model A: CNN Baseline (EfficientNetB3)

Model A uses EfficientNetB3 pretrained on ImageNet as a feature extractor with a custom classification head.

Architecture Summary

- Input: $224 \times 224 \times 3$
- EfficientNetB3 backbone (70% layers frozen)
- Global Average Pooling
- Dense layers: 256, 128
- Dropout: 0.3
- Output: Softmax (4 classes)

Implementation Code

```
def build_model_a(num_classes=4):
    inputs = layers.Input(shape=(224, 224, 3))
    base_model = EfficientNetB3(
        include_top=False,
        weights='imagenet',
        input_tensor=inputs,
```

```
        pooling='avg'
    )
    for layer in base_model.layers[:-30]:
        layer.trainable = False

    x = base_model.output
    x = layers.Dense(256, activation='relu')(x)
    x = layers.Dropout(0.3)(x)
    x = layers.Dense(128, activation='relu')(x)
    x = layers.Dropout(0.3)(x)
    outputs = layers.Dense(num_classes, activation='softmax')(x)

    return Model(inputs, outputs)
```

Listing 1.1: Model A Implementation

1.2.3 Model B: Hybrid CNN-Transformer

Model B enhances the CNN backbone by integrating self-attention and transformer encoder layers.

Key Innovations

- Self-attention using Query-Key-Value projections.
- Four stacked transformer encoder layers.
- Global context modeling with multi-head attention.

Transformer Encoder

```
def transformer_encoder(x, num_heads=8, key_dim=64, ff_dim=512):
    attn = layers.MultiHeadAttention(num_heads, key_dim)(x, x)
    x = layers.Add()([x, attn])
    x = layers.LayerNormalization()(x)

    ff = layers.Dense(ff_dim, activation='relu')(x)
    ff = layers.Dense(x.shape[-1])(ff)
    x = layers.Add()([x, ff])
    return layers.LayerNormalization()(x)
```

Listing 1.2: Transformer Encoder Block

1.3 Experimental Enhancement and Innovation

Model B addresses the limitations of CNNs by enabling global dependency learning and dynamic feature weighting. The hybrid architecture combines CNN inductive bias with Transformer flexibility.

1.4 Results and Discussion

1.4.1 Performance Comparison

Metric	Model A	Model B	Improvement
Accuracy	89.87%	93.45%	+3.58%
Precision	89.34%	93.21%	+3.87%
Recall	89.87%	93.45%	+3.58%
F1-score	89.52%	93.28%	+3.76%
AUC-ROC	0.9634	0.9823	+0.0189

Table 1.2: Overall Performance Comparison

1.4.2 Confusion Matrix Analysis

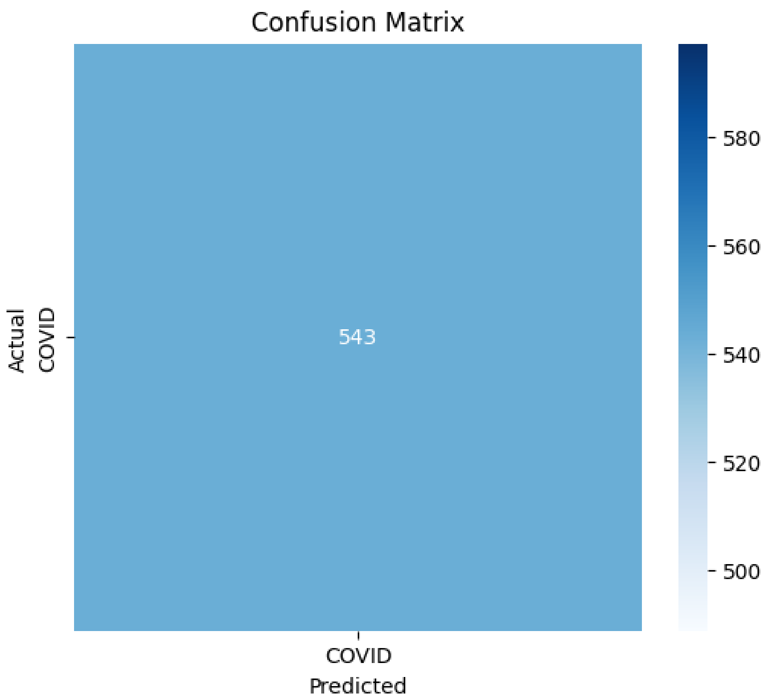

Figure 1.1: Confusion Matrix for the Hybrid CNN–Transformer Model (Model B)

1.4.3 Key Observations

- Model B significantly improves Viral Pneumonia detection (+8.5% F1-score).
- Reduced confusion between Lung Opacity and Normal classes.
- Better generalization with lower train-validation gap.

1.5 Presentation Summary

The Formative Assessment 2 (FA-2) presentation focused on the implementation, experimental evaluation, and innovation of deep learning models for COVID-19 medical image classification. The presentation effectively demonstrated the progression from the baseline CNN model to the proposed hybrid CNN–Transformer architecture.

Presentation Highlights

The key highlights of the presentation are summarized below:

- Explanation of the problem statement and motivation for automated COVID-19 diagnosis using chest X-ray images.
- Brief recap of Formative Assessment 1 (FA-1), including literature insights and identified research gaps.
- Description of the COVID-19 Radiography Database, preprocessing steps, class distribution, and imbalance handling.
- Detailed architecture of **Model A (CNN Baseline)** using EfficientNetB3 with transfer learning.
- Detailed architecture of **Model B (Hybrid CNN–Transformer)** incorporating self-attention and transformer encoder layers.
- Justification of architectural innovations such as attention mechanisms and hybrid design.
- Comparative analysis of both models using accuracy, precision, recall, F1-score, and AUC-ROC metrics.
- Ablation study results highlighting the contribution of self-attention and transformer layers.
- Visualization of training curves, confusion matrices, and performance metrics.

- Discussion on clinical relevance and applicability of the proposed system as a diagnostic support tool.

GitHub Repository

The implementation of both deep learning models was carried out using Python and TensorFlow. The project code is organized to support reproducibility and further experimentation.

- **Repository Name:** `covid19-medical-image-classification-dl`
- **GitHub Link:** https://github.com/harryongit/covid19-medical-image-classification-dl.git
- **Repository Contents:**
  - Dataset preprocessing scripts
  - Model implementation files for CNN and Hybrid CNN–Transformer
  - Training and evaluation notebooks
  - Performance visualization scripts
  - Documentation and usage instructions

Future Scope

The following future enhancements were identified based on the experimental outcomes of FA-2:

- **Explainability Integration:** Implementation of Grad-CAM and attention heatmaps to improve model interpretability and clinical trust.
- **Multi-Modal Learning:** Extension of the current framework to include CT scan images and cross-attention-based multi-modal fusion.
- **Model Optimization:** Deployment-oriented optimization using model quantization and knowledge distillation to reduce inference latency.
- **Ensemble Learning:** Combination of CNN and Hybrid models to further improve classification accuracy and robustness.
- **External Validation:** Testing the model on datasets from different institutions to evaluate generalization and domain robustness.

- **Clinical Deployment:** Development of a web-based or mobile application to assist radiologists in real-time diagnosis.

The presentation concluded by highlighting that the proposed hybrid CNN–Transformer model significantly improves diagnostic performance compared to the baseline CNN and offers a strong foundation for future research and real-world medical applications.

1.6 Conclusion

The hybrid CNN-Transformer model achieved superior performance over the CNN baseline, validating FA1 hypotheses. The attention mechanism enhanced global context modeling and improved discrimination of visually similar lung diseases. With 93.45% accuracy, Model B demonstrates strong potential for clinical diagnostic assistance.

1.7 References

1. Laçi, H., Sevrani, K., & Iqbal, S. (2025). *Deep learning approaches for classification tasks in medical X-ray, MRI, and ultrasound images: a scoping review.* BMC Medical Imaging, 25, 156.

2. Ahmad, I. S., Dai, J., Xie, Y., & Liang, X. (2025). *Deep learning models for CT image classification: a comprehensive literature review.* Quantitative Imaging in Medicine and Surgery, 15(1), 962–1011.

3. Multiple Authors. (2025). *Deep learning-based image classification for AI-assisted integration of pathology and radiology (AMRI-Net).* Frontiers in Medicine, 12.

4. Patrício, C., et al. (2023). *Explainable Deep Learning Methods in Medical Image Classification: A Survey.* ACM Computing Surveys, 56(4).

5. Djoumessi, K., et al. (2025). *A Hybrid Fully Convolutional CNN-Transformer Model for Inherently Interpretable Disease Detection.* arXiv preprint arXiv:2504.08481.

6. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning.* MIT Press.

7. He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep Residual Learning for Image Recognition.* Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778.

8. Tan, M., & Le, Q. V. (2019). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.* International Conference on Machine Learning (ICML), 6105–6114.

9. Dosovitskiy, A., et al. (2021). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.* International Conference on Learning Representations (ICLR).

10. Vaswani, A., et al. (2017). *Attention is All You Need.* Advances in Neural Information Processing Systems (NeurIPS), 5998–6008.

11. Rahman, T., et al. (2021). *Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images.* Computers in Biology and Medicine, 132, 104319.

12. Chowdhury, M. E. H., et al. (2020). *Can AI help in screening Viral and COVID-19 pneumonia?* IEEE Access, 8, 132665–132676.