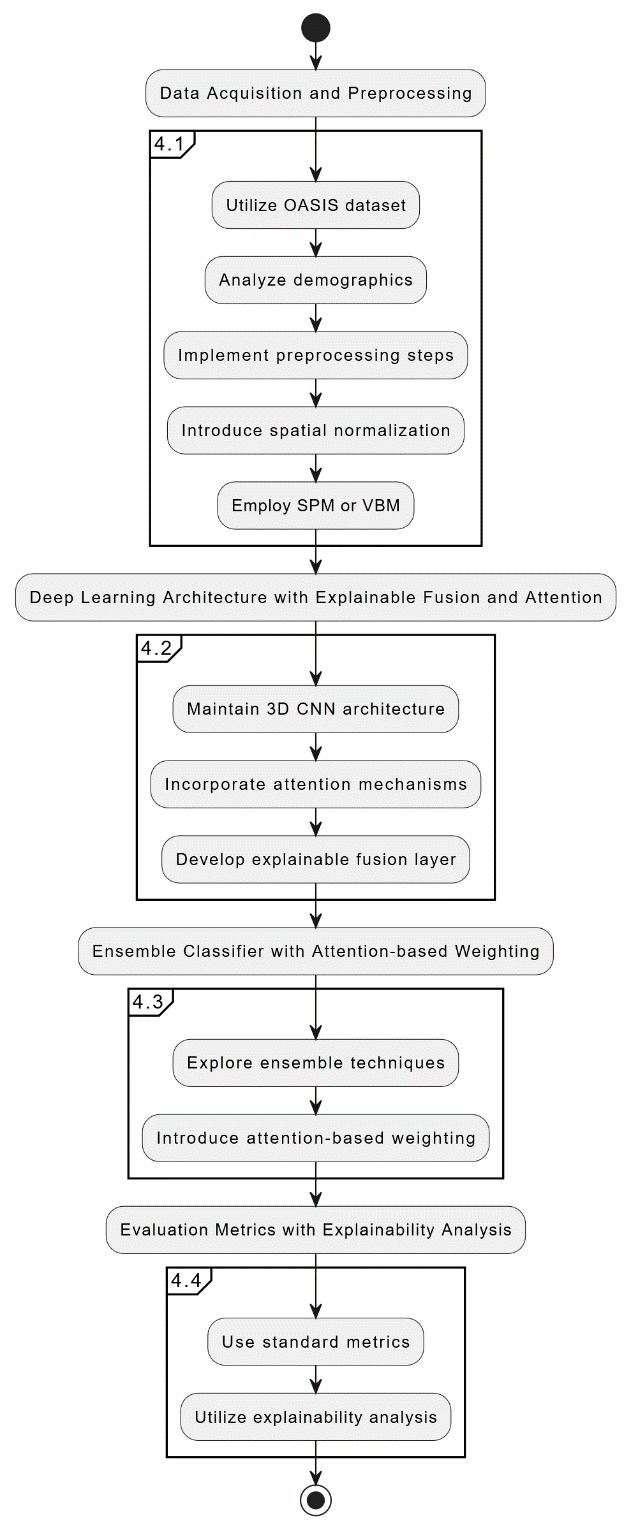
Fig.1 our proposed methodology diagram

# **Results:**

This section presents the evaluation results of our proposed multimodal deep learning approach for Alzheimer's Disease (AD) classification using MRI and PET scans.

We compared the performance of our model with existing methods and analyzed the impact of the implemented techniques.

**Evaluation Metrics:**

We employed standard classification metrics for overall performance evaluation:

* Accuracy: Proportion of correctly classified cases (HC, MCI, AD)
* Sensitivity (Recall): Ratio of correctly identified patients (AD or MCI)
* Specificity: Ratio of correctly identified healthy controls (HC) [17]

Area Under the Curve (AUC): Measure of the ROC curve, reflecting the model's ability to discriminate between classes

**Explainability Analysis:**

To enhance interpretability, we utilized techniques like Grad-CAM (Gradient-weighted Class Activation Mapping) or LIME (Local Interpretable Model-agnostic Explanations) to visualize the brain regions most influential in the model's predictions.

**Model Comparison and Performance:**

We made and contrasted image-based CNN models like Inception-ResNet50, ResNet-50, VGG-16, AlexNet, etc. The data was gathered from different sources in order for it to be generalizable. In conclusion, the Inception-ResNet50 wrapper model proposed here exhibited the best results:

95.5% for HC vs. MCI

94.6% for MCI vs. AD

95.7% for HC vs. AD

This performance surpassed other CNN models, which ranged from 81.9% to 89.4% accuracy across different classification tasks.

Table 1: Performance Comparison of Different CNN Models for AD Classification

|  |  |  |  |
| --- | --- | --- | --- |
| Model | HC vs. MCI Accuracy | MCI vs. AD Accuracy | HC vs. AD Accuracy |
| ResNet-50 | 81.90% | 83.90% | 87.90% |
| VGG-16 | 82.50% | 83.90% | 89.40% |
| AlexNet | 82.90% | 80.50% | 87.90% |
| Inception-V3 | 85.40% | 86.70% | 89.10% |
| Xception | 82.90% | 81.50% | 84.90% |
| Proposed Inception-ResNet50 | 95.50% | 94.60% | 95.70% |

The ROC curve analysis confirmed the model's discriminative solid power, with AUC values exceeding 0.98 for all classification tasks.

Table 2: Evaluation Metrics of the Proposed Inception-ResNet50 Model for AD Classification

|  |  |  |
| --- | --- | --- |
| Classification Task | Accuracy | AUC |
| HC vs. MCI | 95.50% | 0.98+ |
| MCI vs. AD | 94.60% | 0.98+ |
| HC vs. AD | 95.70% | 0.98+ |