## **Approach to Classify AD vs. CN**

The paper presents a novel hybrid deep learning model for classifying **Alzheimer’s Disease (AD)** versus **Cognitively Normal (CN)** individuals using **T1-weighted structural MRI (sMRI)** slices. The model is designed to efficiently capture both **local brain abnormalities** (like hippocampal atrophy) and **global structural changes** associated with AD, while maintaining low computational cost.

### **1. Hybrid Architecture Design**

The core innovation lies in **alternating two building blocks** across four stages of the network:

* **MBConv + Block Attention**
* **PConv + Grid Attention**

This alternating design enables the model to handle both **fine-grained details** and **long-range dependencies** within sMRI images.

### **2. MBConv + Block Attention (Capturing Local Features)**

* **MBConv** is a lightweight convolution block (from MobileNet) that uses depthwise separable convolutions and squeeze-excitation modules to extract spatial and channel-wise features efficiently.
* **Block Attention** divides the feature map into local patches (P×P blocks) and applies **relative multi-head self-attention (Rel-MSA)** to learn relationships within each block.

This module is effective in focusing on **local regions of interest**, such as the hippocampus or ventricles—areas commonly affected in AD.

### **3. PConv + Grid Attention (Capturing Global Context)**

* **Partial Convolution (PConv)** operates only on a subset of the channels, reducing redundancy and computational load.
* **Grid Attention** splits the input into large grid regions (G×G) and applies global self-attention to model **interactions between distant brain regions**.

This allows the network to understand **distributed atrophy patterns** seen across the brain in AD.

### **4. Inverted Residual Feed Forward Network (IRFFN)**

Instead of a standard Transformer MLP block, the model uses an **IRFFN**, which combines:

* Depthwise convolutions,
* Residual (skip) connections,
* Pointwise (1×1) convolutions.

This module helps preserve spatial structure and improves feature fusion from previous layers.

### **5. Classification and Training**

* After passing through the attention and convolution blocks, a **Global Average Pooling** and **1×1 convolution** reduce the feature dimensions and mix channel information.
* A final **softmax classifier** predicts whether the input sMRI slice is from an AD or CN subject.
* The model is trained and validated on the **ADNI dataset** using **10-fold cross-validation**, with data augmentation (random flips and shifts) to improve generalization.

### **6. Performance and Interpretation**

* The model achieves **97.29% accuracy** and **97.14% AUC** in the AD vs. CN task.
* Using **Grad-CAM**, the authors show that the model reliably attends to known AD-related regions (e.g., hippocampus, parietal and temporal lobes), demonstrating both high performance and interpretability.

## **Limitations**

### **1. Reliance on 2D Slices**

The model processes 2D sMRI slices rather than full 3D volumes. While this simplifies training and reduces computation, it **loses critical spatial context across adjacent slices**, which could result in false positives or inconsistent predictions. Brain atrophy in AD often follows 3D patterns, and modeling the full volumetric data could improve accuracy and robustness.

### **2. Limited Generalizability Across Datasets**

The model is trained and evaluated exclusively on the **ADNI dataset**, which follows a consistent imaging protocol. Its **generalizability to other datasets or clinical environments** (with different scanners, populations, or noise levels) remains untested. This poses a potential challenge for real-world deployment unless external validation is conducted.