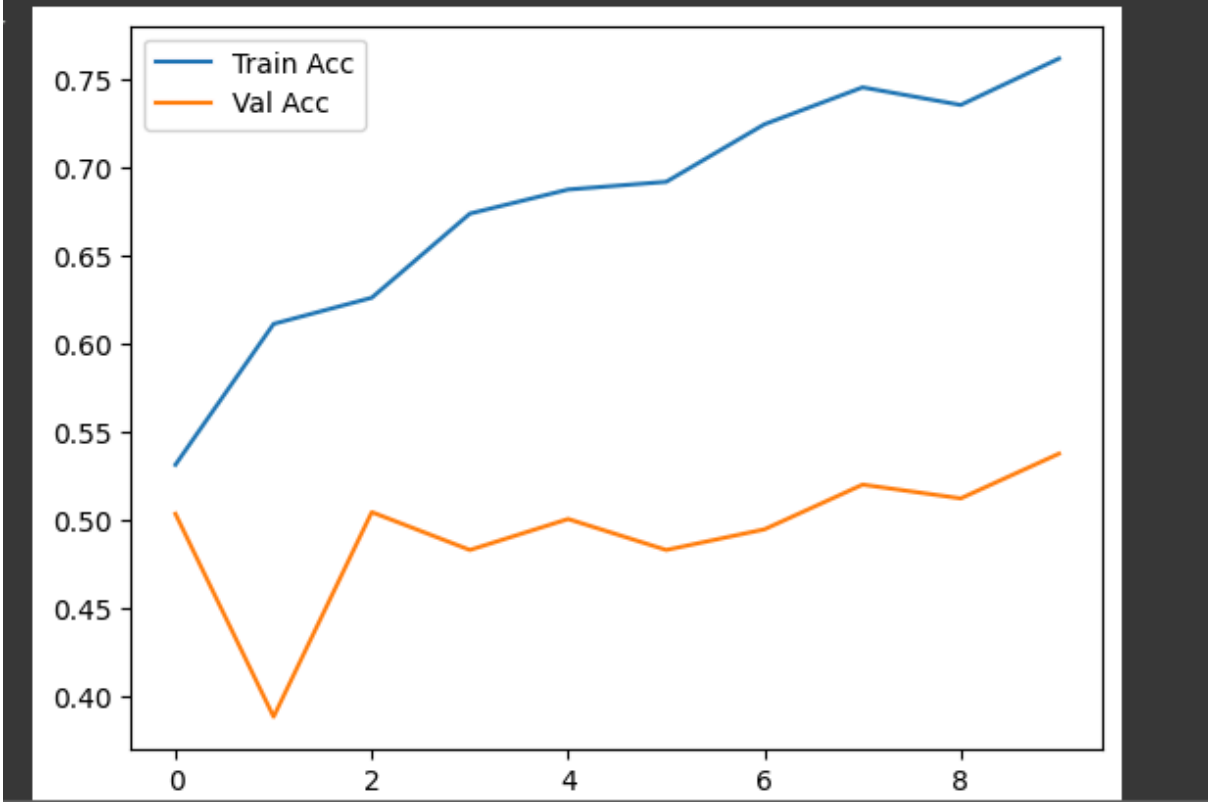
Without Any FineTunning

```
loss, acc = model.evaluate(test_generator)
print(f"Test Accuracy: {acc:.2f}")
```

```
2/2 ━━━━━━━━━━━━━━━━━━━━ 5s 4s/step - accuracy: 0.5333 - loss: 1.2454
Test Accuracy: 0.55
```

```
plt.show()
```



Observations

1. Overfitting:

Model is learning the training set well but not generalizing. This is a classic case when we freeze all convolutional layers and have a small fully connected head.

2. Validation Accuracy Plateau:

Indicates that the model cannot extract more relevant features with the frozen base alone.

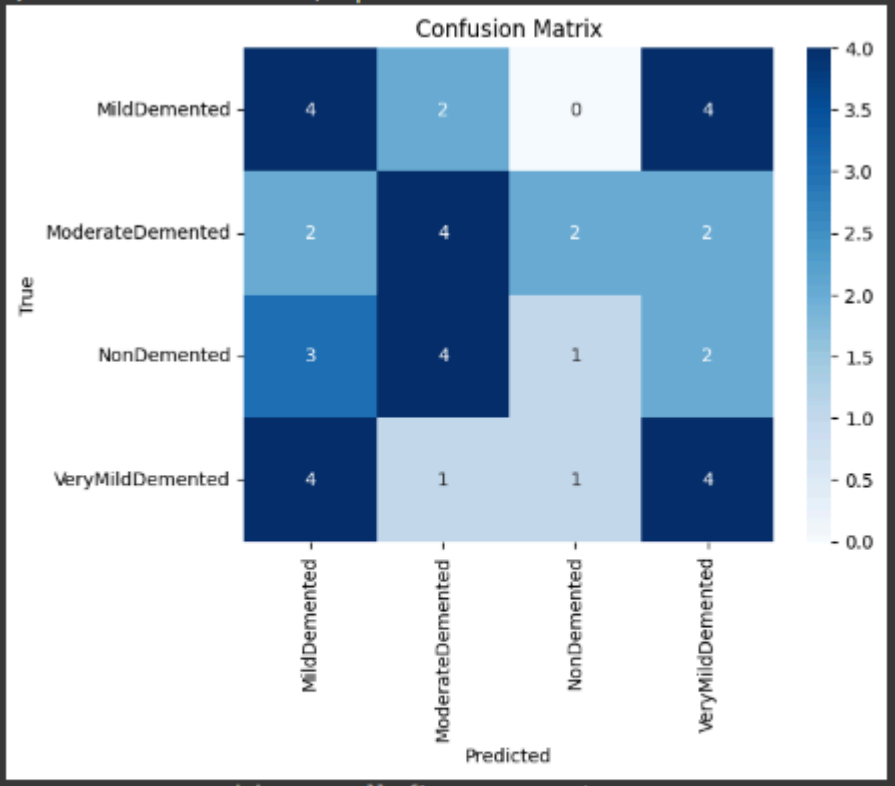
3. Small Dataset + Class Imbalance:

The Alzheimer's dataset (especially “ModerateDemented”) is imbalanced — this hurts generalization.

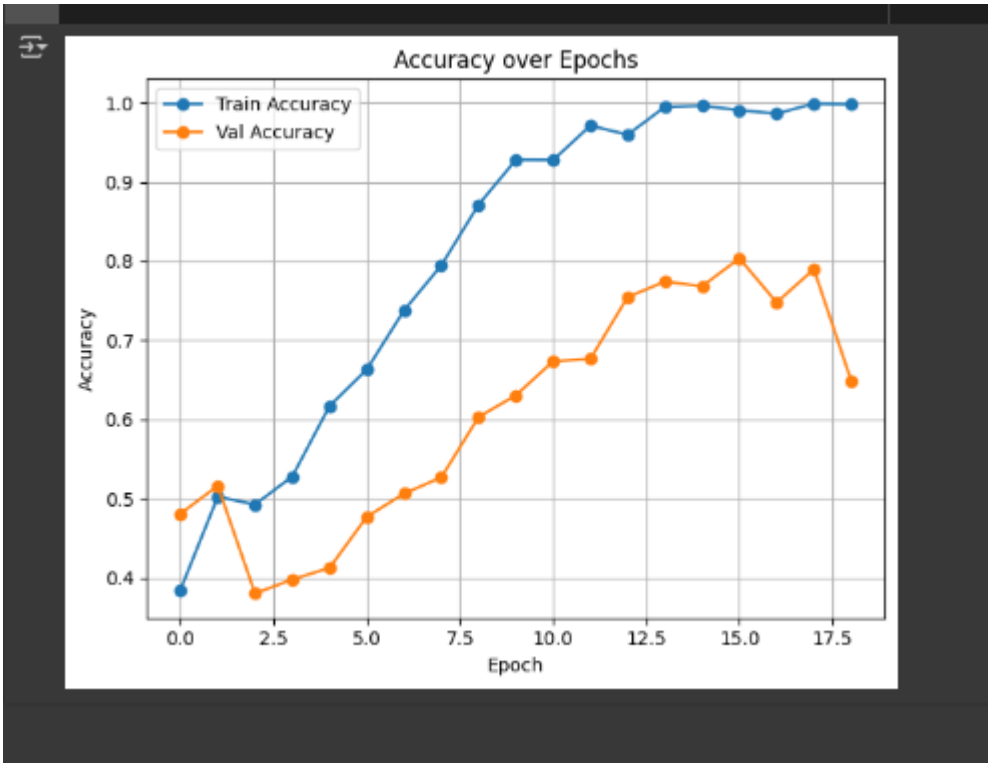
After Fine-Tuning, Accuracy = 80%

```
Found 40 images belonging to 4 classes.
2/2 ━━━━━━━━━━━━━━━━━━━━ 5s 4s/step - accuracy: 0.7937 - loss: 1.2028
Test Accuracy: 80.00%
```

```
2/2 ━━━━━━━━━━━━━━━━━━━━ 4s 1s/step
```



	precision	recall	f1-score	support
MildDemented	0.31	0.40	0.35	10
ModerateDemented	0.36	0.40	0.38	10
NonDemented	0.25	0.10	0.14	10
VeryMildDemented	0.33	0.40	0.36	10
accuracy			0.33	40
macro avg	0.31	0.33	0.31	40
weighted avg	0.31	0.33	0.31	40


Accuracy over Epochs
Train Accuracy
Val Accuracy
Accuracy
Epoch
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5