Model_Evaluation

Rank	Model	Training Accuracy	Validation Accuracy	Test Accuracy	Training Time (sec)	Inference Time (sec)	Deployable	Comment
1	Softmax 3	0.9949	0.981	0.9976	325.79	0.0612	Yes	Best Choice - Strong test accuracy, high validation accuracy (not too perfect), fastest inference, lowest training time.
2	Sigmoid 3	0.9942	0.9714	0.9905	529.02	0.1257	Yes	Good alternative- High accuracy but slightly slower inference than Softmax 3. Slightly lower validation accuracy suggests mild overfitting tendency.
3	MobileNetV2 3	0.999	0.9833	0.9941	1505.6	0.1174	No	Overkill - Good accuracy but takes too long to train and is much slower in inference. Some signs of overfitting, but further validation needed.
4	Softmax 2	0.9912	0.9786	0.9953	450.78	0.1173	Yes	Slightly weaker Softmax - High efficiency, but validation accuracy is lower, suggesting a small risk of overfitting.
5	Sigmoid 2	0.9949	0.9881	0.9988	416.68	0.1458	Yes	Decent choice - Fast and accurate, but inference is slower than Softmax 3. If inference speed is not an issue, it's a reasonable alternative.
6	MobileNetV2 2	0.9986	0.9952	1	1493.31	0.1205	No	Possible Overfitting - Test accuracy is suspiciously high (1.0000), validation is slightly lower. Requires real-world validation.
7	MobileNetV2 1	0.9997	1	1	1484.98	0.1658	No	Likely Overfitting - Training and test accuracy are both perfect. A model this good is too good to be true. Caution advised.
8	Softmax 1	0.9912	0.9952	1	571.14	0.1119	No	Unrealistically perfect - Perfect test accuracy and validation accuracy. Overfitting is a strong concern. Inference time is not the best either.
9	Sigmoid 1	0.9959	0.9952	1	605.94	0.1129	No	Similar to Softmax 1 - Possible overfitting risk. Training time is too long compared to its marginal improvement in performance.

Summary of Model Selection Decision

1. Selecting the Best Model for Deployment

After analyzing multiple models based on accuracy, efficiency, and overfitting tendencies, Softmax 3 is the best choice for deployment.

- Strong test accuracy (0.9976) and high validation accuracy (0.9810) ensure good generalization.
- Fastest inference time (0.0612 sec/sample) makes it the most efficient for real-world use.
- Lowest training time (325.79 sec) minimizes resource consumption.
- Unlike some models, it does not show signs of extreme overfitting, making it reliable for deployment.

Alternative options:

- Sigmoid 3 is a reasonable alternative but has slightly slower inference and lower validation accuracy.
- MobileNetV2 models were not selected due to significantly higher computational costs without a major accuracy advantage.

2. Addressing the Python-Generated "No Overfitting" Statement

The Python-generated summary stated "No Overfitting" for all models. However, a deeper individual analysis suggests:

- Some models (e.g., MobileNetV2 1, Softmax 1, Sigmoid 1) had perfect test accuracy (1.0000), which is suspiciously high.
- Models where test accuracy is higher than validation accuracy may indicate memorization rather than true generalization.
- Extreme accuracy should be validated further, as real-world performance may differ.

Conclusion:

- While Python's summary did not explicitly detect overfitting, individual model analysis suggests potential overfitting risks in certain models (hence, they were excluded).
- Softmax 3 is selected because it strikes the best balance between accuracy, efficiency, and generalization.

3. Final Decision & Deployment Note

Softmax 3 is the most practical model for deployment, offering the best trade-off between accuracy and efficiency.
Models with perfect test accuracy (1.0000) require real-world validation before being considered reliable.
The final selection prioritizes real-world applicability over theoretical accuracy.