

	Optimising Accuracy	Optimising Latency
transforms used in data augmentation	<p>Rescaling: Normalizes pixel values to 0-1 range.</p> <p>Rotation: Randomly rotates images up to <math>\pm 1</math> degree.</p> <p>Zoom: Randomly zooms images up to <math>\pm 10\%</math>.</p> <p>Width Shift: Randomly shifts images horizontally up to <math>\pm 10\%</math> of width.</p> <p>Height Shift: Randomly shifts images vertically up to <math>\pm 10\%</math> of height.</p> <p>Shear: Randomly applies shearing transformation up to <math>\pm 10\%</math>.</p> <p>Horizontal Flip: Randomly flips images horizontally.</p> <p>Fill Mode: Fills new pixels with the nearest copied pixel.</p>	<p>Rescaling: Normalizes pixel values to 0-1 range.</p> <p>Rotation: Randomly rotates images up to <math>\pm 1</math> degree.</p> <p>Zoom: Randomly zooms images up to <math>\pm 10\%</math>.</p> <p>Width Shift: Randomly shifts images horizontally up to <math>\pm 10\%</math> of width.</p> <p>Height Shift: Randomly shifts images vertically up to <math>\pm 10\%</math> of height.</p> <p>Shear: Randomly applies shearing transformation up to <math>\pm 10\%</math>.</p> <p>Horizontal Flip: Randomly flips images horizontally.</p> <p>Fill Mode: Fills new pixels with the nearest copied pixel.</p>
base model	Exception, 88MB, 79%, 109.4ms	MobileNetV2, 14MB, 71.3%, 25.9ms
number of epochs, optimizer, and learning rate used to train classification head	20 epochs, Adam @ 0.001	10 epochs Adam @ 0.001
number of layers un-frozen	120	100
number of epochs, optimizer, and learning rate used to further fine-tune the model	20 epochs, Adam @0.00001	5 epochs, Adam @0.0001
final accuracy on evaluation set (test set)	0.8775	0.8503

### MobileNetV2:

Classification Head Training (10 epochs, Adam @ 0.001): Shorter duration prevents overfitting, striking a balance between training adequately and avoiding prolonged time.

Further Fine-Tuning (5 epochs, Adam @ 0.0001): Short duration refines parameters without risking overfitting due to lower complexity.

### Xception:

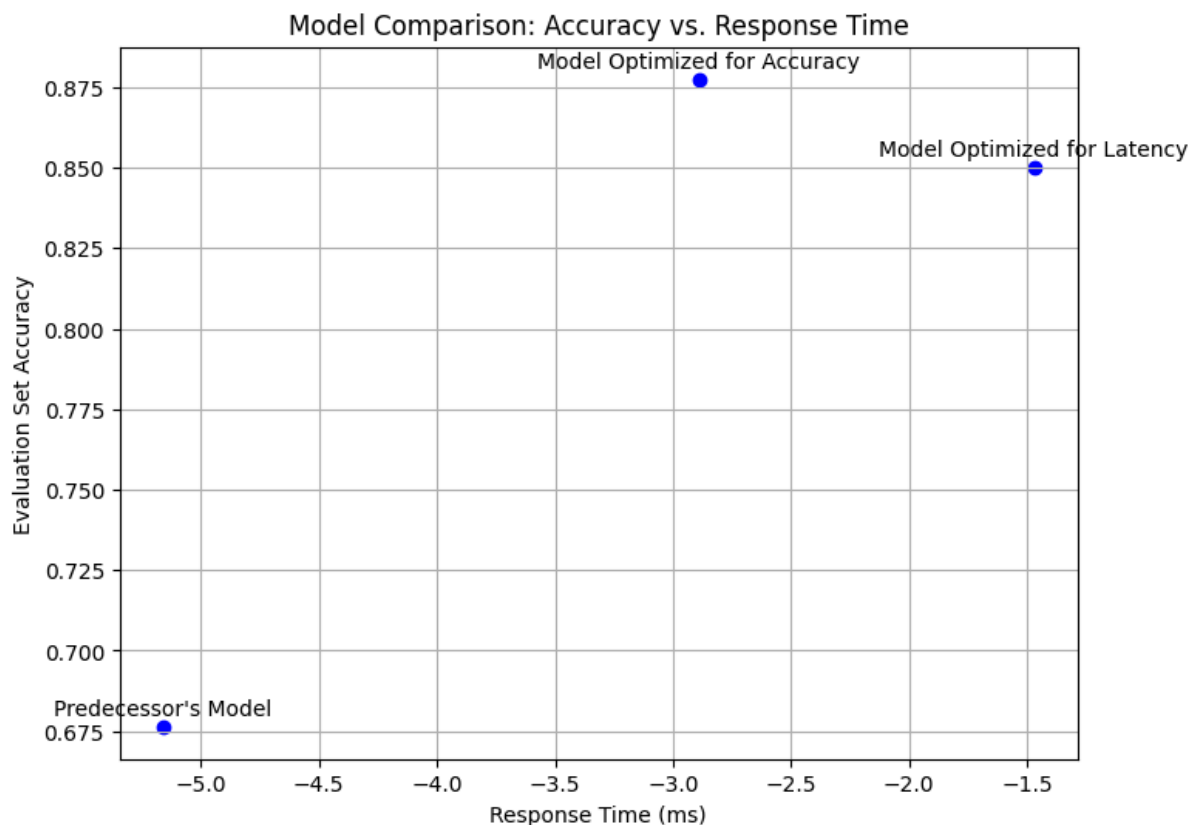
Classification Head Training (20 epochs, Adam @ 0.001): Longer duration allows capturing intricate features, providing ample time for convergence.

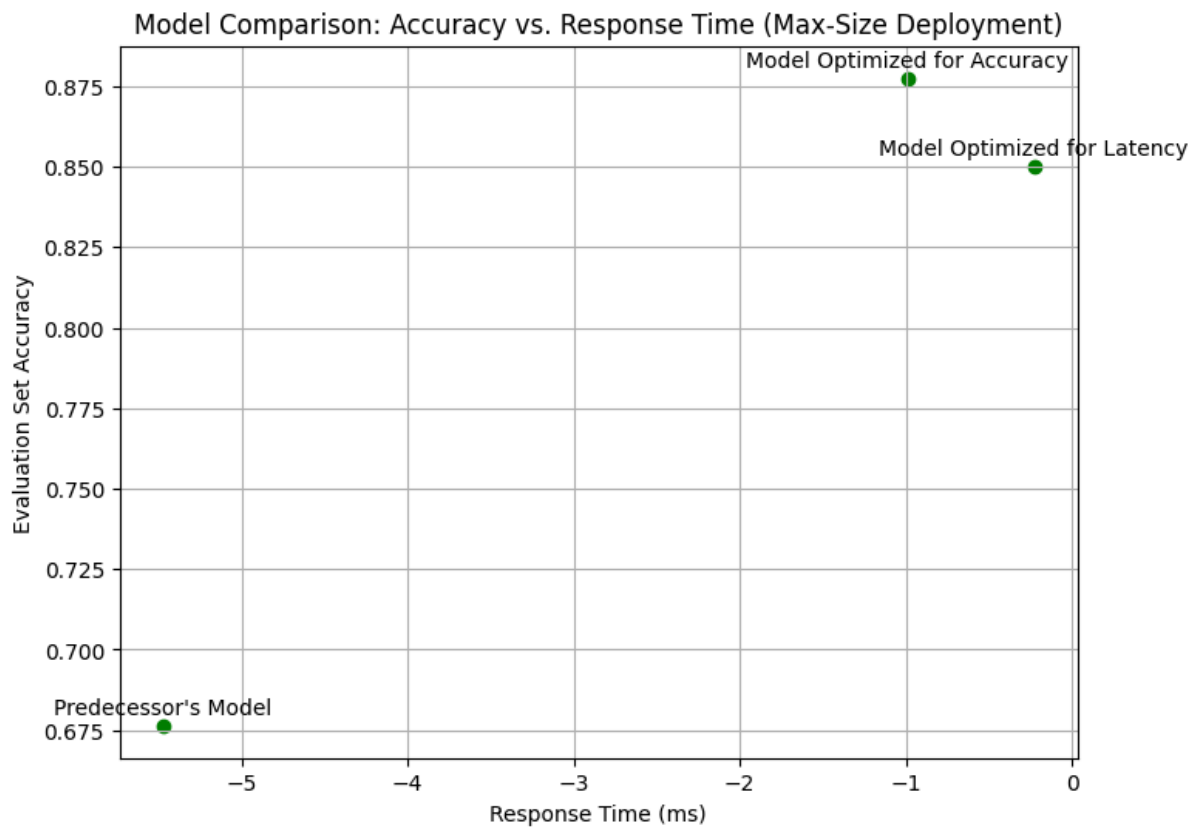
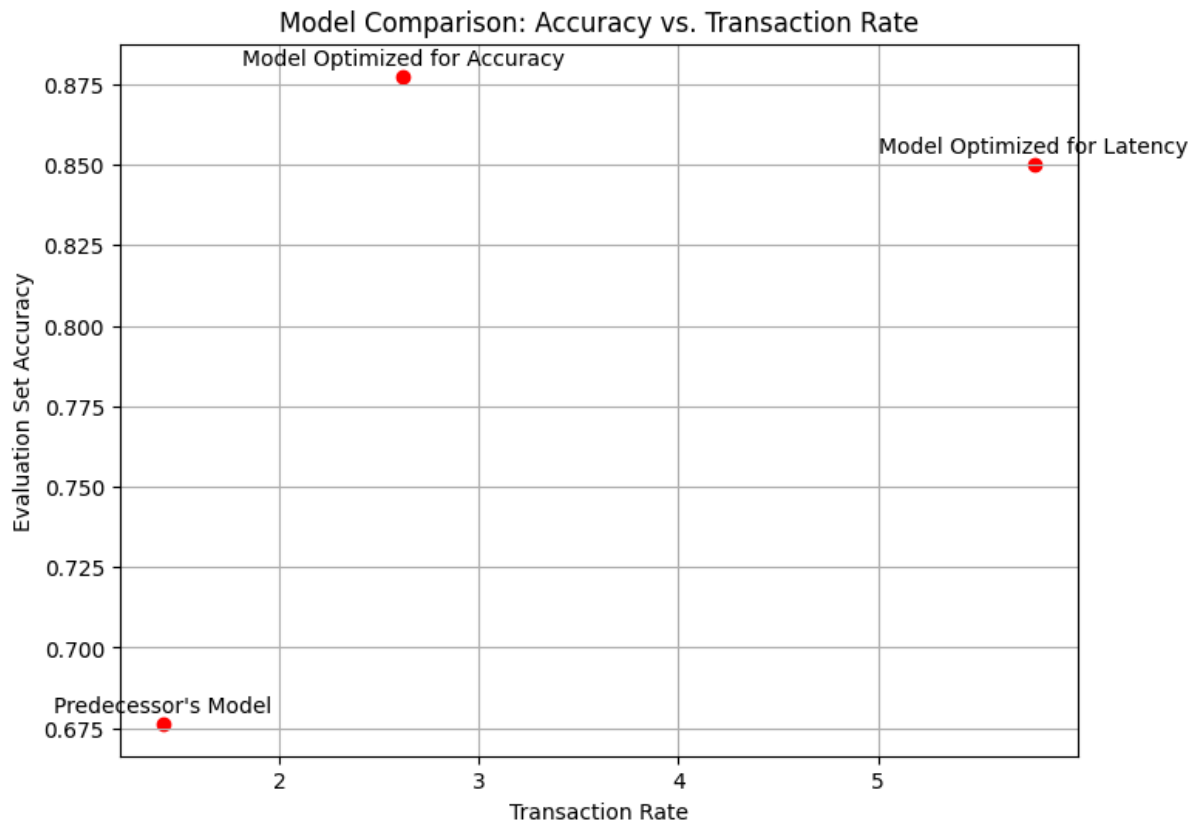
Further Fine-Tuning (20 epochs, Adam @ 0.00001): Ensures full adaptation to task-specific data while minimizing overfitting risk with ultra-low learning rate.

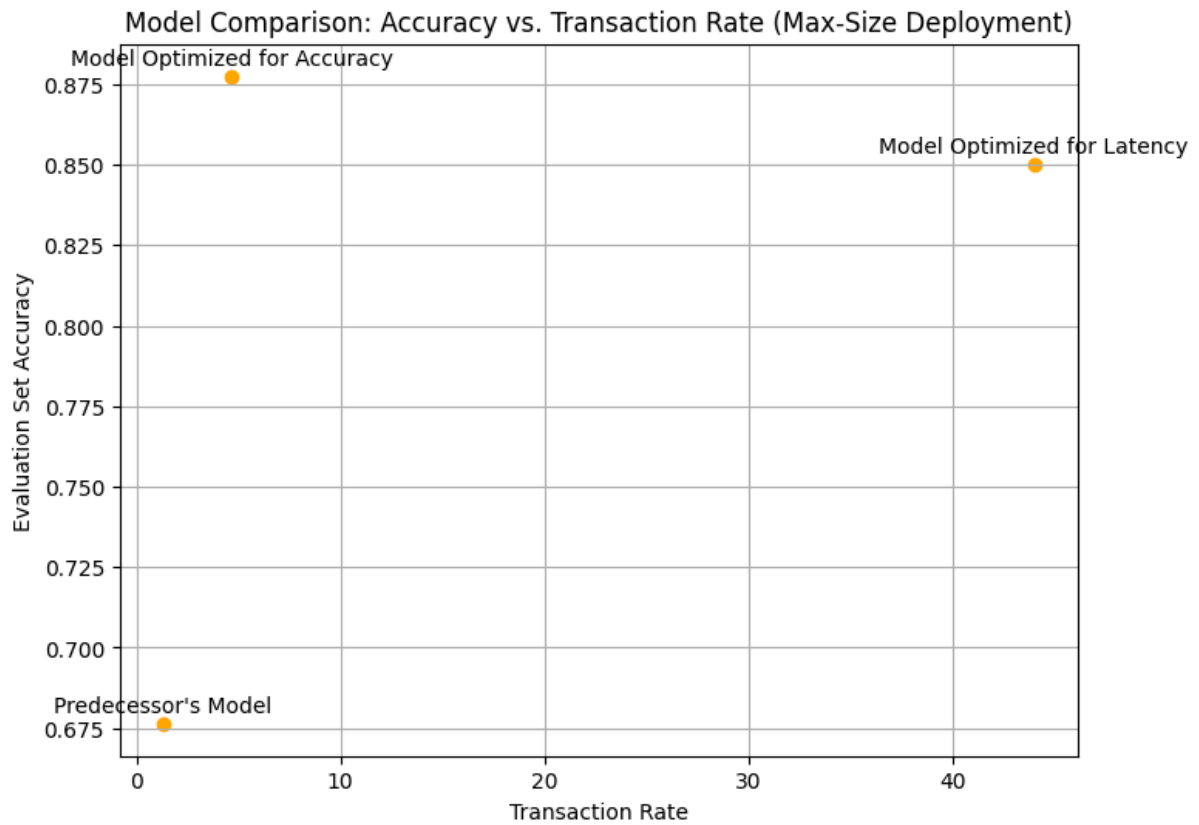
### Reasoning:

MobileNetV2: Moderate initial learning rate stabilizes convergence, while smaller fine-tuning rate aids gradual parameter adjustment.

Xception: Similar initial learning rate ensures effective learning; ultra-low fine-tuning rate facilitates slow parameter adjustment, retaining learned features.







	Previous	Accuracy	Latency
Number of replicas	19	19	19
CPU resource requests	1405m (35%)	3405m (85%)	3405m (85%)
Memory resource requests	2529882Ki (31%)	6724186Ki (83%)	6724186Ki (83%)
CPU resource limits	3400m (85%)	7400m (185%)	7400m (185%)
Memory resource limits	5470397696 (66%)	14060332288 (170%)	14060332288 (170%)

	Accuracy	Latency
minReplicas	1	1
maxReplicas	10	8
targetCPUUtilizationPercentage	40	80
CPU resource requests	0.5	0.5
Memory resource requests	1Gi	1Gi
CPU resource limits	3	3
Memory resource limits	5	5

