

Device in use: cuda

=== Advanced CNN with Attention Mechanisms for CIFAR-10 Classification ===

=== Experiment 1: Baseline Model (Enhanced from Claude_v2) ===

[Epoch: 1/75] Loss: 2.0344, Accuracy: 27.74%, Learning rate: 0.000400, Time: 18.40s
[Epoch: 2/75] Loss: 1.8320, Accuracy: 40.20%, Learning rate: 0.000400, Time: 18.42s
[Epoch: 3/75] Loss: 1.7085, Accuracy: 46.83%, Learning rate: 0.000400, Time: 18.63s
[Epoch: 4/75] Loss: 1.6439, Accuracy: 50.52%, Learning rate: 0.000400, Time: 18.62s
[Epoch: 5/75] Loss: 1.5512, Accuracy: 55.24%, Learning rate: 0.000400, Time: 18.72s
[Epoch: 6/75] Loss: 1.4851, Accuracy: 58.29%, Learning rate: 0.000400, Time: 18.77s
[Epoch: 7/75] Loss: 1.4764, Accuracy: 59.57%, Learning rate: 0.000400, Time: 18.90s
[Epoch: 8/75] Loss: 1.4539, Accuracy: 60.99%, Learning rate: 0.000400, Time: 18.91s
[Epoch: 9/75] Loss: 1.4105, Accuracy: 62.56%, Learning rate: 0.000400, Time: 18.93s
[Epoch: 10/75] Loss: 1.3966, Accuracy: 63.32%, Learning rate: 0.000400, Time: 18.94s
[Epoch: 11/75] Loss: 1.3394, Accuracy: 66.28%, Learning rate: 0.000400, Time: 18.90s
[Epoch: 12/75] Loss: 1.3471, Accuracy: 65.97%, Learning rate: 0.000400, Time: 18.94s
[Epoch: 13/75] Loss: 1.3151, Accuracy: 66.82%, Learning rate: 0.000400, Time: 18.91s
[Epoch: 14/75] Loss: 1.3187, Accuracy: 67.42%, Learning rate: 0.000400, Time: 18.99s
[Epoch: 15/75] Loss: 1.2849, Accuracy: 68.62%, Learning rate: 0.000400, Time: 19.00s
[Epoch: 16/75] Loss: 1.2786, Accuracy: 68.79%, Learning rate: 0.000400, Time: 18.96s
[Epoch: 17/75] Loss: 1.2788, Accuracy: 68.70%, Learning rate: 0.000400, Time: 18.97s
[Epoch: 18/75] Loss: 1.2602, Accuracy: 69.65%, Learning rate: 0.000400, Time: 18.99s
[Epoch: 19/75] Loss: 1.2499, Accuracy: 69.56%, Learning rate: 0.000400, Time: 18.96s

[Epoch: 20/75] Loss: 1.2327, Accuracy: 70.84%, Learning rate: 0.000400,
Time: 18.95s
[Epoch: 21/75] Loss: 1.2291, Accuracy: 70.96%, Learning rate: 0.000400,
Time: 18.88s
[Epoch: 22/75] Loss: 1.2315, Accuracy: 71.02%, Learning rate: 0.000400,
Time: 18.90s
[Epoch: 23/75] Loss: 1.2097, Accuracy: 71.55%, Learning rate: 0.000400,
Time: 18.95s
[Epoch: 24/75] Loss: 1.1966, Accuracy: 72.47%, Learning rate: 0.000400,
Time: 18.90s
[Epoch: 25/75] Loss: 1.2196, Accuracy: 71.51%, Learning rate: 0.000400,
Time: 18.90s
[Epoch: 26/75] Loss: 1.2080, Accuracy: 72.01%, Learning rate: 0.000400,
Time: 18.85s
[Epoch: 27/75] Loss: 1.2003, Accuracy: 72.18%, Learning rate: 0.000400,
Time: 18.89s
[Epoch: 28/75] Loss: 1.1714, Accuracy: 73.37%, Learning rate: 0.000400,
Time: 18.87s
[Epoch: 29/75] Loss: 1.1894, Accuracy: 72.86%, Learning rate: 0.000400,
Time: 18.90s
[Epoch: 30/75] Loss: 1.2044, Accuracy: 71.77%, Learning rate: 0.000400,
Time: 18.88s
[Epoch: 31/75] Loss: 1.1374, Accuracy: 75.21%, Learning rate: 0.000400,
Time: 18.92s
[Epoch: 32/75] Loss: 1.1526, Accuracy: 73.83%, Learning rate: 0.000400,
Time: 18.85s
[Epoch: 33/75] Loss: 1.1341, Accuracy: 74.18%, Learning rate: 0.000400,
Time: 18.90s
[Epoch: 34/75] Loss: 1.1557, Accuracy: 73.69%, Learning rate: 0.000400,
Time: 18.94s
[Epoch: 35/75] Loss: 1.1267, Accuracy: 74.95%, Learning rate: 0.000400,
Time: 18.87s
[Epoch: 36/75] Loss: 1.1572, Accuracy: 73.76%, Learning rate: 0.000400,
Time: 18.86s
[Epoch: 37/75] Loss: 1.1001, Accuracy: 76.39%, Learning rate: 0.000400,
Time: 18.80s
[Epoch: 38/75] Loss: 1.1120, Accuracy: 75.11%, Learning rate: 0.000400,
Time: 18.90s
[Epoch: 39/75] Loss: 1.1320, Accuracy: 74.72%, Learning rate: 0.000400,
Time: 18.83s
[Epoch: 40/75] Loss: 1.0905, Accuracy: 75.61%, Learning rate: 0.000400,
Time: 18.80s

[Epoch: 41/75] Loss: 1.0721, Accuracy: 76.80%, Learning rate: 0.000400,
Time: 18.81s
[Epoch: 42/75] Loss: 1.1061, Accuracy: 76.28%, Learning rate: 0.000400,
Time: 18.86s
[Epoch: 43/75] Loss: 1.1164, Accuracy: 75.29%, Learning rate: 0.000400,
Time: 18.84s
[Epoch: 44/75] Loss: 1.0946, Accuracy: 76.31%, Learning rate: 0.000400,
Time: 18.90s
[Epoch: 45/75] Loss: 1.1064, Accuracy: 75.79%, Learning rate: 0.000400,
Time: 18.88s
[Epoch: 46/75] Loss: 1.0822, Accuracy: 77.02%, Learning rate: 0.000400,
Time: 18.81s
[Epoch: 47/75] Loss: 1.0622, Accuracy: 77.73%, Learning rate: 0.000400,
Time: 18.81s
[Epoch: 48/75] Loss: 1.0833, Accuracy: 77.49%, Learning rate: 0.000400,
Time: 18.85s
[Epoch: 49/75] Loss: 1.0649, Accuracy: 77.12%, Learning rate: 0.000400,
Time: 18.91s
[Epoch: 50/75] Loss: 1.0811, Accuracy: 76.96%, Learning rate: 0.000400,
Time: 18.81s
[Epoch: 51/75] Loss: 1.0662, Accuracy: 77.87%, Learning rate: 0.000400,
Time: 18.86s
[Epoch: 52/75] Loss: 1.1002, Accuracy: 75.52%, Learning rate: 0.000400,
Time: 18.83s
[Epoch: 53/75] Loss: 1.0653, Accuracy: 77.43%, Learning rate: 0.000400,
Time: 18.79s
[Epoch: 54/75] Loss: 1.0621, Accuracy: 77.64%, Learning rate: 0.000400,
Time: 18.81s
[Epoch: 55/75] Loss: 1.0765, Accuracy: 76.87%, Learning rate: 0.000400,
Time: 18.82s
[Epoch: 56/75] Loss: 1.0451, Accuracy: 77.34%, Learning rate: 0.000400,
Time: 18.80s
[Epoch: 57/75] Loss: 1.0426, Accuracy: 78.45%, Learning rate: 0.000400,
Time: 18.78s
[Epoch: 58/75] Loss: 1.0759, Accuracy: 77.20%, Learning rate: 0.000400,
Time: 18.80s
[Epoch: 59/75] Loss: 1.0592, Accuracy: 77.81%, Learning rate: 0.000400,
Time: 18.81s
[Epoch: 60/75] Loss: 1.0635, Accuracy: 77.47%, Learning rate: 0.000400,
Time: 18.78s
[Epoch: 61/75] Loss: 1.0413, Accuracy: 78.23%, Learning rate: 0.000400,
Time: 18.77s

[Epoch: 62/75] Loss: 1.0498, Accuracy: 77.58%, Learning rate: 0.000400,
Time: 18.82s
[Epoch: 63/75] Loss: 1.0174, Accuracy: 79.40%, Learning rate: 0.000400,
Time: 18.81s
[Epoch: 64/75] Loss: 1.0367, Accuracy: 78.46%, Learning rate: 0.000400,
Time: 18.80s
[Epoch: 65/75] Loss: 1.0472, Accuracy: 77.64%, Learning rate: 0.000400,
Time: 18.77s
[Epoch: 66/75] Loss: 1.0168, Accuracy: 79.35%, Learning rate: 0.000400,
Time: 18.81s
[Epoch: 67/75] Loss: 1.0385, Accuracy: 78.39%, Learning rate: 0.000400,
Time: 18.79s
[Epoch: 68/75] Loss: 1.0230, Accuracy: 78.62%, Learning rate: 0.000400,
Time: 18.84s
[Epoch: 69/75] Loss: 1.0328, Accuracy: 78.16%, Learning rate: 0.000400,
Time: 18.73s
[Epoch: 70/75] Loss: 1.0439, Accuracy: 77.59%, Learning rate: 0.000400,
Time: 18.78s
[Epoch: 71/75] Loss: 1.0265, Accuracy: 78.76%, Learning rate: 0.000400,
Time: 18.82s
[Epoch: 72/75] Loss: 1.0187, Accuracy: 79.05%, Learning rate: 0.000400,
Time: 18.79s
[Epoch: 73/75] Loss: 1.0179, Accuracy: 78.68%, Learning rate: 0.000400,
Time: 18.76s
[Epoch: 74/75] Loss: 0.9952, Accuracy: 79.57%, Learning rate: 0.000400,
Time: 18.77s
[Epoch: 75/75] Loss: 0.9921, Accuracy: 80.32%, Learning rate: 0.000400,
Time: 18.79s
Test set accuracy: 92.09%

=== Experiment 2: SE-Net Model (Adding Squeeze-and-Excitation) ===

[Epoch: 1/75] Loss: 2.0859, Accuracy: 25.35%, Learning rate: 0.001000,
Time: 21.02s
[Epoch: 2/75] Loss: 1.9074, Accuracy: 36.39%, Learning rate: 0.000998,
Time: 20.94s
[Epoch: 3/75] Loss: 1.8150, Accuracy: 42.75%, Learning rate: 0.000996,
Time: 20.95s
[Epoch: 4/75] Loss: 1.7535, Accuracy: 46.77%, Learning rate: 0.000993,
Time: 21.01s
[Epoch: 5/75] Loss: 1.6799, Accuracy: 49.90%, Learning rate: 0.000989,
Time: 20.93s
[Epoch: 6/75] Loss: 1.6071, Accuracy: 53.01%, Learning rate: 0.000984,
Time: 20.94s

[Epoch: 7/75] Loss: 1.5958, Accuracy: 54.79%, Learning rate: 0.000979,
Time: 21.00s
[Epoch: 8/75] Loss: 1.5520, Accuracy: 56.96%, Learning rate: 0.000972,
Time: 20.88s
[Epoch: 9/75] Loss: 1.5478, Accuracy: 58.26%, Learning rate: 0.000965,
Time: 20.92s
[Epoch: 10/75] Loss: 1.4998, Accuracy: 59.56%, Learning rate: 0.000957,
Time: 20.96s
[Epoch: 11/75] Loss: 1.5069, Accuracy: 59.55%, Learning rate: 0.000948,
Time: 20.97s
[Epoch: 12/75] Loss: 1.4716, Accuracy: 61.39%, Learning rate: 0.000938,
Time: 20.90s
[Epoch: 13/75] Loss: 1.4047, Accuracy: 64.33%, Learning rate: 0.000928,
Time: 20.93s
[Epoch: 14/75] Loss: 1.4196, Accuracy: 64.11%, Learning rate: 0.000916,
Time: 20.91s
[Epoch: 15/75] Loss: 1.4420, Accuracy: 63.66%, Learning rate: 0.000905,
Time: 20.97s
[Epoch: 16/75] Loss: 1.3952, Accuracy: 65.56%, Learning rate: 0.000892,
Time: 21.02s
[Epoch: 17/75] Loss: 1.3785, Accuracy: 65.43%, Learning rate: 0.000878,
Time: 20.92s
[Epoch: 18/75] Loss: 1.3463, Accuracy: 66.58%, Learning rate: 0.000864,
Time: 20.97s
[Epoch: 19/75] Loss: 1.3277, Accuracy: 67.90%, Learning rate: 0.000850,
Time: 20.90s
[Epoch: 20/75] Loss: 1.3271, Accuracy: 67.90%, Learning rate: 0.000835,
Time: 20.88s
[Epoch: 21/75] Loss: 1.3161, Accuracy: 68.96%, Learning rate: 0.000819,
Time: 20.92s
[Epoch: 22/75] Loss: 1.3057, Accuracy: 69.03%, Learning rate: 0.000802,
Time: 20.95s
[Epoch: 23/75] Loss: 1.3048, Accuracy: 68.82%, Learning rate: 0.000785,
Time: 21.01s
[Epoch: 24/75] Loss: 1.3232, Accuracy: 68.34%, Learning rate: 0.000768,
Time: 20.96s
[Epoch: 25/75] Loss: 1.2850, Accuracy: 69.82%, Learning rate: 0.000750,
Time: 20.94s
[Epoch: 26/75] Loss: 1.3074, Accuracy: 68.21%, Learning rate: 0.000732,
Time: 20.91s
[Epoch: 27/75] Loss: 1.2952, Accuracy: 70.12%, Learning rate: 0.000713,
Time: 20.96s

[Epoch: 28/75] Loss: 1.2628, Accuracy: 71.43%, Learning rate: 0.000694, Time: 21.00s
[Epoch: 29/75] Loss: 1.2474, Accuracy: 71.19%, Learning rate: 0.000674, Time: 20.97s
[Epoch: 30/75] Loss: 1.2345, Accuracy: 72.21%, Learning rate: 0.000655, Time: 20.86s
[Epoch: 31/75] Loss: 1.2300, Accuracy: 72.29%, Learning rate: 0.000634, Time: 20.99s
[Epoch: 32/75] Loss: 1.2504, Accuracy: 71.00%, Learning rate: 0.000614, Time: 20.91s
[Epoch: 33/75] Loss: 1.2333, Accuracy: 72.53%, Learning rate: 0.000594, Time: 20.87s
[Epoch: 34/75] Loss: 1.2301, Accuracy: 71.92%, Learning rate: 0.000573, Time: 20.94s
[Epoch: 35/75] Loss: 1.2204, Accuracy: 72.84%, Learning rate: 0.000552, Time: 20.97s
[Epoch: 36/75] Loss: 1.1674, Accuracy: 75.01%, Learning rate: 0.000531, Time: 20.93s
[Epoch: 37/75] Loss: 1.1945, Accuracy: 73.41%, Learning rate: 0.000510, Time: 20.92s
[Epoch: 38/75] Loss: 1.2190, Accuracy: 73.26%, Learning rate: 0.000490, Time: 20.98s
[Epoch: 39/75] Loss: 1.2035, Accuracy: 73.54%, Learning rate: 0.000469, Time: 20.89s
[Epoch: 40/75] Loss: 1.1881, Accuracy: 74.03%, Learning rate: 0.000448, Time: 20.96s
[Epoch: 41/75] Loss: 1.1652, Accuracy: 74.92%, Learning rate: 0.000427, Time: 20.98s
[Epoch: 42/75] Loss: 1.1769, Accuracy: 74.25%, Learning rate: 0.000406, Time: 21.02s
[Epoch: 43/75] Loss: 1.1641, Accuracy: 75.52%, Learning rate: 0.000386, Time: 20.91s
[Epoch: 44/75] Loss: 1.1548, Accuracy: 76.15%, Learning rate: 0.000366, Time: 20.95s
[Epoch: 45/75] Loss: 1.1449, Accuracy: 75.70%, Learning rate: 0.000345, Time: 20.93s
[Epoch: 46/75] Loss: 1.1841, Accuracy: 74.24%, Learning rate: 0.000326, Time: 20.90s
[Epoch: 47/75] Loss: 1.1503, Accuracy: 75.54%, Learning rate: 0.000306, Time: 20.95s
[Epoch: 48/75] Loss: 1.1678, Accuracy: 75.50%, Learning rate: 0.000287, Time: 20.94s

[Epoch: 49/75] Loss: 1.1667, Accuracy: 75.22%, Learning rate: 0.000268,
Time: 20.87s
[Epoch: 50/75] Loss: 1.1127, Accuracy: 76.58%, Learning rate: 0.000250,
Time: 20.87s
[Epoch: 51/75] Loss: 1.1206, Accuracy: 76.87%, Learning rate: 0.000232,
Time: 20.92s
[Epoch: 52/75] Loss: 1.0828, Accuracy: 77.67%, Learning rate: 0.000215,
Time: 20.93s
[Epoch: 53/75] Loss: 1.1157, Accuracy: 77.41%, Learning rate: 0.000198,
Time: 20.88s
[Epoch: 54/75] Loss: 1.1017, Accuracy: 77.82%, Learning rate: 0.000181,
Time: 20.94s
[Epoch: 55/75] Loss: 1.1093, Accuracy: 76.68%, Learning rate: 0.000165,
Time: 20.94s
[Epoch: 56/75] Loss: 1.1421, Accuracy: 75.95%, Learning rate: 0.000150,
Time: 20.90s
[Epoch: 57/75] Loss: 1.1442, Accuracy: 75.48%, Learning rate: 0.000136,
Time: 20.89s
[Epoch: 58/75] Loss: 1.1266, Accuracy: 76.93%, Learning rate: 0.000122,
Time: 20.91s
[Epoch: 59/75] Loss: 1.1180, Accuracy: 76.51%, Learning rate: 0.000108,
Time: 20.88s
[Epoch: 60/75] Loss: 1.1045, Accuracy: 77.76%, Learning rate: 0.000095,
Time: 20.89s
[Epoch: 61/75] Loss: 1.1212, Accuracy: 77.21%, Learning rate: 0.000084,
Time: 20.92s
[Epoch: 62/75] Loss: 1.1174, Accuracy: 76.79%, Learning rate: 0.000072,
Time: 20.91s
[Epoch: 63/75] Loss: 1.0937, Accuracy: 78.01%, Learning rate: 0.000062,
Time: 20.86s
[Epoch: 64/75] Loss: 1.0857, Accuracy: 77.36%, Learning rate: 0.000052,
Time: 20.89s
[Epoch: 65/75] Loss: 1.1035, Accuracy: 77.21%, Learning rate: 0.000043,
Time: 20.92s
[Epoch: 66/75] Loss: 1.0894, Accuracy: 78.10%, Learning rate: 0.000035,
Time: 20.83s
[Epoch: 67/75] Loss: 1.1140, Accuracy: 77.86%, Learning rate: 0.000028,
Time: 20.95s
[Epoch: 68/75] Loss: 1.0862, Accuracy: 78.59%, Learning rate: 0.000021,
Time: 20.95s
[Epoch: 69/75] Loss: 1.0998, Accuracy: 77.76%, Learning rate: 0.000016,
Time: 20.88s

[Epoch: 70/75] Loss: 1.0723, Accuracy: 78.85%, Learning rate: 0.000011,
Time: 20.88s
[Epoch: 71/75] Loss: 1.1313, Accuracy: 76.16%, Learning rate: 0.000007,
Time: 20.90s
[Epoch: 72/75] Loss: 1.0911, Accuracy: 78.61%, Learning rate: 0.000004,
Time: 20.94s
[Epoch: 73/75] Loss: 1.0679, Accuracy: 78.92%, Learning rate: 0.000002,
Time: 20.83s
[Epoch: 74/75] Loss: 1.0840, Accuracy: 77.42%, Learning rate: 0.000000,
Time: 21.02s
[Epoch: 75/75] Loss: 1.0709, Accuracy: 79.56%, Learning rate: 0.000000,
Time: 20.92s
Test set accuracy: 93.16%

=== Experiment 3: ECA-Net Model (Efficient Channel Attention) ===

[Epoch: 1/75] Loss: 2.0760, Accuracy: 25.35%, Learning rate: 0.000001,
Time: 20.93s
[Epoch: 2/75] Loss: 1.9624, Accuracy: 32.44%, Learning rate: 0.000994,
Time: 20.81s
[Epoch: 3/75] Loss: 1.8995, Accuracy: 36.80%, Learning rate: 0.000976,
Time: 20.72s
[Epoch: 4/75] Loss: 1.7862, Accuracy: 43.53%, Learning rate: 0.000946,
Time: 20.79s
[Epoch: 5/75] Loss: 1.7057, Accuracy: 48.15%, Learning rate: 0.000905,
Time: 20.81s
[Epoch: 6/75] Loss: 1.6434, Accuracy: 51.50%, Learning rate: 0.000854,
Time: 20.82s
[Epoch: 7/75] Loss: 1.5835, Accuracy: 55.14%, Learning rate: 0.000794,
Time: 20.74s
[Epoch: 8/75] Loss: 1.5623, Accuracy: 56.41%, Learning rate: 0.000998,
Time: 20.75s
[Epoch: 9/75] Loss: 1.5196, Accuracy: 58.46%, Learning rate: 0.000984,
Time: 20.76s
[Epoch: 10/75] Loss: 1.4925, Accuracy: 60.31%, Learning rate: 0.000958,
Time: 20.67s
[Epoch: 11/75] Loss: 1.4347, Accuracy: 62.37%, Learning rate: 0.000001,
Time: 20.84s
[Epoch: 12/75] Loss: 1.4101, Accuracy: 64.16%, Learning rate: 0.000825,
Time: 20.75s
[Epoch: 13/75] Loss: 1.4380, Accuracy: 63.31%, Learning rate: 0.000794,
Time: 20.81s
[Epoch: 14/75] Loss: 1.3954, Accuracy: 65.26%, Learning rate: 0.001000,
Time: 20.78s

[Epoch: 15/75] Loss: 1.4042, Accuracy: 64.72%, Learning rate: 0.000998,
Time: 20.84s
[Epoch: 16/75] Loss: 1.3983, Accuracy: 64.94%, Learning rate: 0.000992,
Time: 20.85s
[Epoch: 17/75] Loss: 1.3424, Accuracy: 67.10%, Learning rate: 0.000984,
Time: 20.79s
[Epoch: 18/75] Loss: 1.3529, Accuracy: 67.46%, Learning rate: 0.000972,
Time: 20.71s
[Epoch: 19/75] Loss: 1.3346, Accuracy: 68.03%, Learning rate: 0.000958,
Time: 20.79s
[Epoch: 20/75] Loss: 1.3175, Accuracy: 68.93%, Learning rate: 0.000941,
Time: 20.75s
[Epoch: 21/75] Loss: 1.2725, Accuracy: 70.60%, Learning rate: 0.000921,
Time: 20.83s
[Epoch: 22/75] Loss: 1.2938, Accuracy: 69.83%, Learning rate: 0.000898,
Time: 20.73s
[Epoch: 23/75] Loss: 1.2774, Accuracy: 71.00%, Learning rate: 0.000873,
Time: 20.80s
[Epoch: 24/75] Loss: 1.2690, Accuracy: 71.08%, Learning rate: 0.000846,
Time: 20.79s
[Epoch: 25/75] Loss: 1.2587, Accuracy: 71.82%, Learning rate: 0.000817,
Time: 20.76s
[Epoch: 26/75] Loss: 1.2620, Accuracy: 71.63%, Learning rate: 0.000786,
Time: 20.75s
[Epoch: 27/75] Loss: 1.2531, Accuracy: 72.41%, Learning rate: 0.001000,
Time: 20.78s
[Epoch: 28/75] Loss: 1.2631, Accuracy: 71.68%, Learning rate: 0.000996,
Time: 20.82s
[Epoch: 29/75] Loss: 1.2446, Accuracy: 72.54%, Learning rate: 0.000990,
Time: 20.71s
[Epoch: 30/75] Loss: 1.2364, Accuracy: 72.61%, Learning rate: 0.000981,
Time: 20.74s
[Epoch: 31/75] Loss: 1.2320, Accuracy: 72.58%, Learning rate: 0.000969,
Time: 20.81s
[Epoch: 32/75] Loss: 1.2186, Accuracy: 73.63%, Learning rate: 0.000954,
Time: 20.75s
[Epoch: 33/75] Loss: 1.2109, Accuracy: 73.31%, Learning rate: 0.000936,
Time: 20.74s
[Epoch: 34/75] Loss: 1.1876, Accuracy: 74.22%, Learning rate: 0.000915,
Time: 20.71s
[Epoch: 35/75] Loss: 1.1656, Accuracy: 75.68%, Learning rate: 0.000892,
Time: 20.80s

[Epoch: 36/75] Loss: 1.1954, Accuracy: 74.70%, Learning rate: 0.000866,
Time: 20.80s
[Epoch: 37/75] Loss: 1.1597, Accuracy: 75.60%, Learning rate: 0.000839,
Time: 20.77s
[Epoch: 38/75] Loss: 1.1755, Accuracy: 75.38%, Learning rate: 0.000809,
Time: 20.73s
[Epoch: 39/75] Loss: 1.1775, Accuracy: 75.65%, Learning rate: 0.000777,
Time: 20.68s
[Epoch: 40/75] Loss: 1.1296, Accuracy: 77.20%, Learning rate: 0.000999,
Time: 20.80s
[Epoch: 41/75] Loss: 1.1494, Accuracy: 76.76%, Learning rate: 0.000995,
Time: 20.78s
[Epoch: 42/75] Loss: 1.1356, Accuracy: 76.78%, Learning rate: 0.000988,
Time: 20.81s
[Epoch: 43/75] Loss: 1.1363, Accuracy: 76.49%, Learning rate: 0.000978,
Time: 20.80s
[Epoch: 44/75] Loss: 1.1449, Accuracy: 76.82%, Learning rate: 0.000965,
Time: 20.78s
[Epoch: 45/75] Loss: 1.1442, Accuracy: 77.03%, Learning rate: 0.000949,
Time: 20.76s
[Epoch: 46/75] Loss: 1.1371, Accuracy: 76.84%, Learning rate: 0.000930,
Time: 20.83s
[Epoch: 47/75] Loss: 1.1396, Accuracy: 77.41%, Learning rate: 0.000909,
Time: 20.79s
[Epoch: 48/75] Loss: 1.1165, Accuracy: 77.18%, Learning rate: 0.000885,
Time: 20.78s
[Epoch: 49/75] Loss: 1.1232, Accuracy: 77.24%, Learning rate: 0.000859,
Time: 20.74s
[Epoch: 50/75] Loss: 1.1071, Accuracy: 78.12%, Learning rate: 0.000831,
Time: 20.72s
[Epoch: 51/75] Loss: 1.1198, Accuracy: 78.03%, Learning rate: 0.000001,
Time: 20.71s
[Epoch: 52/75] Loss: 1.1119, Accuracy: 78.27%, Learning rate: 0.001000,
Time: 20.77s
[Epoch: 53/75] Loss: 1.1217, Accuracy: 77.46%, Learning rate: 0.001000,
Time: 20.72s
[Epoch: 54/75] Loss: 1.1235, Accuracy: 77.84%, Learning rate: 0.000998,
Time: 20.75s
[Epoch: 55/75] Loss: 1.1098, Accuracy: 77.85%, Learning rate: 0.000996,
Time: 20.73s
[Epoch: 56/75] Loss: 1.1045, Accuracy: 78.35%, Learning rate: 0.000994,
Time: 20.77s

[Epoch: 57/75] Loss: 1.0869, Accuracy: 79.01%, Learning rate: 0.000990,
Time: 20.69s
[Epoch: 58/75] Loss: 1.0929, Accuracy: 79.29%, Learning rate: 0.000986,
Time: 20.75s
[Epoch: 59/75] Loss: 1.0841, Accuracy: 79.25%, Learning rate: 0.000981,
Time: 20.73s
[Epoch: 60/75] Loss: 1.0766, Accuracy: 79.19%, Learning rate: 0.000975,
Time: 20.81s
[Epoch: 61/75] Loss: 1.0854, Accuracy: 79.41%, Learning rate: 0.000969,
Time: 20.66s
[Epoch: 62/75] Loss: 1.0929, Accuracy: 79.05%, Learning rate: 0.000961,
Time: 20.67s
[Epoch: 63/75] Loss: 1.0684, Accuracy: 79.58%, Learning rate: 0.000954,
Time: 20.75s
[Epoch: 64/75] Loss: 1.0550, Accuracy: 80.77%, Learning rate: 0.000945,
Time: 20.76s
[Epoch: 65/75] Loss: 1.0410, Accuracy: 80.70%, Learning rate: 0.000936,
Time: 20.78s
[Epoch: 66/75] Loss: 1.0954, Accuracy: 79.38%, Learning rate: 0.000926,
Time: 20.74s
[Epoch: 67/75] Loss: 1.0490, Accuracy: 80.72%, Learning rate: 0.000915,
Time: 20.67s
[Epoch: 68/75] Loss: 1.0739, Accuracy: 80.02%, Learning rate: 0.000904,
Time: 20.86s
[Epoch: 69/75] Loss: 1.0567, Accuracy: 80.24%, Learning rate: 0.000892,
Time: 20.76s
[Epoch: 70/75] Loss: 1.0699, Accuracy: 79.53%, Learning rate: 0.000879,
Time: 20.69s
[Epoch: 71/75] Loss: 1.0759, Accuracy: 80.08%, Learning rate: 0.000866,
Time: 20.82s
[Epoch: 72/75] Loss: 1.0311, Accuracy: 81.09%, Learning rate: 0.000853,
Time: 20.78s
[Epoch: 73/75] Loss: 0.9840, Accuracy: 83.03%, Learning rate: 0.000839,
Time: 20.69s
[Epoch: 74/75] Loss: 0.9976, Accuracy: 82.82%, Learning rate: 0.000824,
Time: 20.69s
[Epoch: 75/75] Loss: 1.0356, Accuracy: 80.98%, Learning rate: 0.000809,
Time: 20.72s
Test set accuracy: 91.64%

=== Experiment 4: Advanced Model with SWA (Stochastic Weight Averaging) ===
[Epoch: 1/75] Loss: 2.1204, Accuracy: 23.25%, Learning rate: 0.001000,
Time: 24.88s

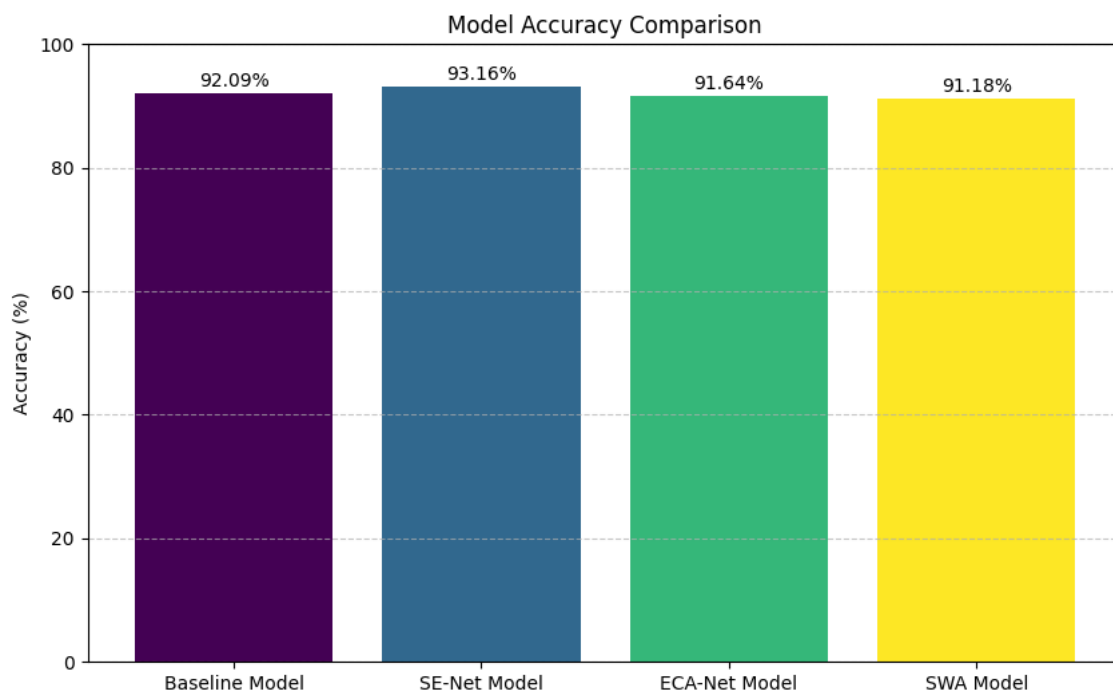
[Epoch: 2/75] Loss: 1.9938, Accuracy: 31.36%, Learning rate: 0.000998,
Time: 24.60s
[Epoch: 3/75] Loss: 1.9204, Accuracy: 36.41%, Learning rate: 0.000996,
Time: 24.75s
[Epoch: 4/75] Loss: 1.8351, Accuracy: 41.01%, Learning rate: 0.000993,
Time: 24.64s
[Epoch: 5/75] Loss: 1.8052, Accuracy: 43.12%, Learning rate: 0.000989,
Time: 24.79s
[Epoch: 6/75] Loss: 1.7554, Accuracy: 46.01%, Learning rate: 0.000984,
Time: 24.61s
[Epoch: 7/75] Loss: 1.7124, Accuracy: 48.28%, Learning rate: 0.000979,
Time: 24.67s
[Epoch: 8/75] Loss: 1.6997, Accuracy: 48.72%, Learning rate: 0.000972,
Time: 24.67s
[Epoch: 9/75] Loss: 1.6565, Accuracy: 51.08%, Learning rate: 0.000965,
Time: 24.64s
[Epoch: 10/75] Loss: 1.6553, Accuracy: 51.42%, Learning rate: 0.000957,
Time: 24.62s
[Epoch: 11/75] Loss: 1.6030, Accuracy: 53.57%, Learning rate: 0.000948,
Time: 24.65s
[Epoch: 12/75] Loss: 1.5755, Accuracy: 55.45%, Learning rate: 0.000938,
Time: 24.66s
[Epoch: 13/75] Loss: 1.5677, Accuracy: 56.14%, Learning rate: 0.000928,
Time: 24.59s
[Epoch: 14/75] Loss: 1.5623, Accuracy: 56.59%, Learning rate: 0.000916,
Time: 24.64s
[Epoch: 15/75] Loss: 1.5448, Accuracy: 56.90%, Learning rate: 0.000905,
Time: 24.63s
[Epoch: 16/75] Loss: 1.5402, Accuracy: 57.49%, Learning rate: 0.000892,
Time: 24.76s
[Epoch: 17/75] Loss: 1.5340, Accuracy: 58.50%, Learning rate: 0.000878,
Time: 24.64s
[Epoch: 18/75] Loss: 1.5297, Accuracy: 58.20%, Learning rate: 0.000864,
Time: 24.62s
[Epoch: 19/75] Loss: 1.4848, Accuracy: 60.21%, Learning rate: 0.000850,
Time: 24.65s
[Epoch: 20/75] Loss: 1.5110, Accuracy: 58.36%, Learning rate: 0.000835,
Time: 24.70s
[Epoch: 21/75] Loss: 1.4796, Accuracy: 60.56%, Learning rate: 0.000819,
Time: 24.63s
[Epoch: 22/75] Loss: 1.4865, Accuracy: 60.32%, Learning rate: 0.000802,
Time: 24.56s

[Epoch: 23/75] Loss: 1.4717, Accuracy: 60.81%, Learning rate: 0.000785,
Time: 24.61s
[Epoch: 24/75] Loss: 1.4563, Accuracy: 61.31%, Learning rate: 0.000768,
Time: 24.70s
[Epoch: 25/75] Loss: 1.4280, Accuracy: 62.44%, Learning rate: 0.000750,
Time: 24.64s
[Epoch: 26/75] Loss: 1.4600, Accuracy: 61.82%, Learning rate: 0.000732,
Time: 24.64s
[Epoch: 27/75] Loss: 1.4112, Accuracy: 63.35%, Learning rate: 0.000713,
Time: 24.71s
[Epoch: 28/75] Loss: 1.3866, Accuracy: 64.42%, Learning rate: 0.000694,
Time: 24.67s
[Epoch: 29/75] Loss: 1.4208, Accuracy: 63.21%, Learning rate: 0.000674,
Time: 24.76s
[Epoch: 30/75] Loss: 1.3977, Accuracy: 64.22%, Learning rate: 0.000655,
Time: 24.64s
[Epoch: 31/75] Loss: 1.3765, Accuracy: 65.42%, Learning rate: 0.000634,
Time: 24.65s
[Epoch: 32/75] Loss: 1.3984, Accuracy: 64.77%, Learning rate: 0.000614,
Time: 24.63s
[Epoch: 33/75] Loss: 1.3999, Accuracy: 64.05%, Learning rate: 0.000594,
Time: 24.60s
[Epoch: 34/75] Loss: 1.3716, Accuracy: 65.77%, Learning rate: 0.000573,
Time: 24.69s
[Epoch: 35/75] Loss: 1.3863, Accuracy: 65.28%, Learning rate: 0.000552,
Time: 24.62s
[Epoch: 36/75] Loss: 1.3750, Accuracy: 65.50%, Learning rate: 0.000531,
Time: 24.62s
[Epoch: 37/75] Loss: 1.3652, Accuracy: 65.98%, Learning rate: 0.000510,
Time: 24.60s
[Epoch: 38/75] Loss: 1.3447, Accuracy: 66.13%, Learning rate: 0.000490,
Time: 24.60s
[Epoch: 39/75] Loss: 1.3356, Accuracy: 67.41%, Learning rate: 0.000469,
Time: 24.56s
[Epoch: 40/75] Loss: 1.3431, Accuracy: 66.59%, Learning rate: 0.000448,
Time: 24.56s
[Epoch: 41/75] Loss: 1.3468, Accuracy: 66.60%, Learning rate: 0.000427,
Time: 24.68s
[Epoch: 42/75] Loss: 1.3060, Accuracy: 67.99%, Learning rate: 0.000406,
Time: 24.61s
[Epoch: 43/75] Loss: 1.3394, Accuracy: 67.41%, Learning rate: 0.000386,
Time: 24.59s

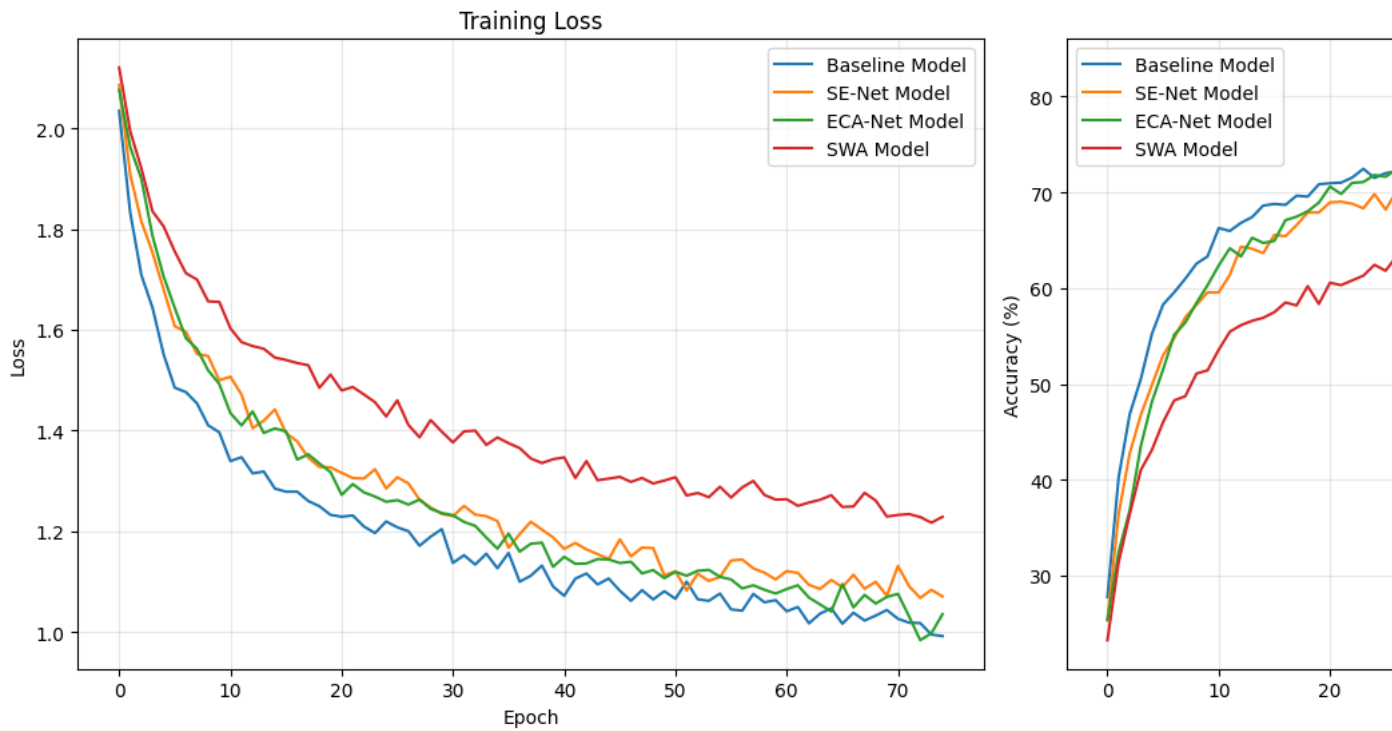
[Epoch: 44/75] Loss: 1.3015, Accuracy: 68.58%, Learning rate: 0.000366,
Time: 24.65s
[Epoch: 45/75] Loss: 1.3046, Accuracy: 68.23%, Learning rate: 0.000345,
Time: 24.59s
[Epoch: 46/75] Loss: 1.3081, Accuracy: 68.48%, Learning rate: 0.000326,
Time: 24.67s
[Epoch: 47/75] Loss: 1.2982, Accuracy: 68.83%, Learning rate: 0.000306,
Time: 24.59s
[Epoch: 48/75] Loss: 1.3059, Accuracy: 68.46%, Learning rate: 0.000287,
Time: 24.67s
[Epoch: 49/75] Loss: 1.2949, Accuracy: 68.50%, Learning rate: 0.000268,
Time: 24.61s
[Epoch: 50/75] Loss: 1.3009, Accuracy: 68.47%, Learning rate: 0.000250,
Time: 24.67s
[Epoch: 51/75] Loss: 1.3073, Accuracy: 67.98%, Learning rate: 0.000232,
Time: 24.62s
[Epoch: 52/75] Loss: 1.2711, Accuracy: 70.03%, Learning rate: 0.000215,
Time: 24.64s
[Epoch: 53/75] Loss: 1.2762, Accuracy: 68.91%, Learning rate: 0.000198,
Time: 24.64s
[Epoch: 54/75] Loss: 1.2677, Accuracy: 70.13%, Learning rate: 0.000181,
Time: 24.61s
[Epoch: 55/75] Loss: 1.2886, Accuracy: 69.17%, Learning rate: 0.000165,
Time: 24.52s
[Epoch: 56/75] Loss: 1.2671, Accuracy: 70.37%, Learning rate: 0.000150,
Time: 24.60s
[Epoch: 57/75] Loss: 1.2872, Accuracy: 69.59%, Learning rate: 0.000136,
Time: 24.69s
[Epoch: 58/75] Loss: 1.3000, Accuracy: 69.72%, Learning rate: 0.000122,
Time: 24.77s
[Epoch: 59/75] Loss: 1.2722, Accuracy: 70.35%, Learning rate: 0.000108,
Time: 24.60s
[Epoch: 60/75] Loss: 1.2630, Accuracy: 70.23%, Learning rate: 0.000095,
Time: 24.65s
[Epoch: 61/75] Loss: 1.2638, Accuracy: 70.53%, Learning rate: 0.000084,
Time: 24.64s
[Epoch: 62/75] Loss: 1.2506, Accuracy: 70.85%, Learning rate: 0.000072,
Time: 24.65s
[Epoch: 63/75] Loss: 1.2571, Accuracy: 69.98%, Learning rate: 0.000062,
Time: 24.64s
[Epoch: 64/75] Loss: 1.2626, Accuracy: 70.42%, Learning rate: 0.000052,
Time: 24.62s

[Epoch: 65/75] Loss: 1.2717, Accuracy: 70.35%, Learning rate: 0.000043, Time: 24.61s
[Epoch: 66/75] Loss: 1.2484, Accuracy: 70.51%, Learning rate: 0.000035, Time: 24.67s
[Epoch: 67/75] Loss: 1.2494, Accuracy: 71.03%, Learning rate: 0.000028, Time: 24.62s
[Epoch: 68/75] Loss: 1.2764, Accuracy: 69.86%, Learning rate: 0.000021, Time: 24.74s
[Epoch: 69/75] Loss: 1.2611, Accuracy: 70.91%, Learning rate: 0.000016, Time: 24.67s
[Epoch: 70/75] Loss: 1.2294, Accuracy: 71.44%, Learning rate: 0.000011, Time: 24.63s
[Epoch: 71/75] Loss: 1.2325, Accuracy: 71.80%, Learning rate: 0.000007, Time: 24.65s
[Epoch: 72/75] Loss: 1.2342, Accuracy: 72.21%, Learning rate: 0.000004, Time: 24.59s
[Epoch: 73/75] Loss: 1.2285, Accuracy: 71.84%, Learning rate: 0.000002, Time: 24.67s
[Epoch: 74/75] Loss: 1.2172, Accuracy: 72.31%, Learning rate: 0.000000, Time: 24.64s
[Epoch: 75/75] Loss: 1.2288, Accuracy: 71.06%, Learning rate: 0.000000, Time: 24.58s
Test set accuracy: 91.18%

=== Result Comparison and Visualization ===

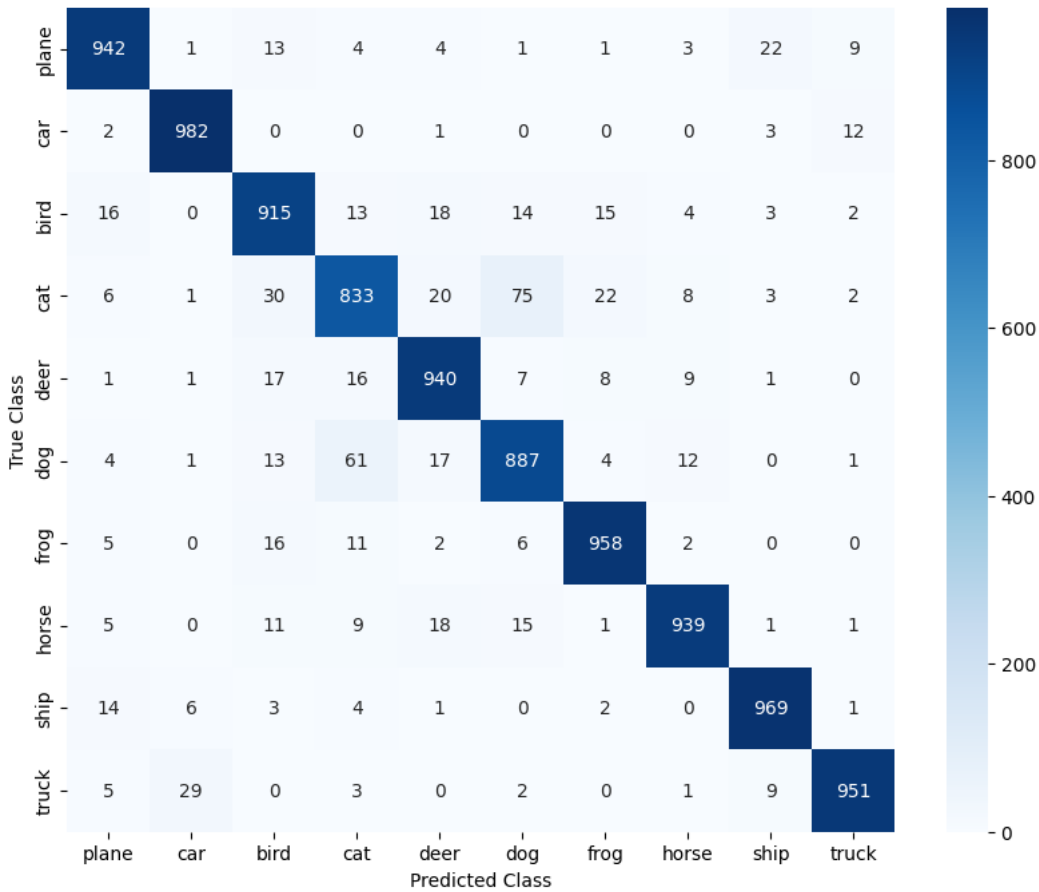


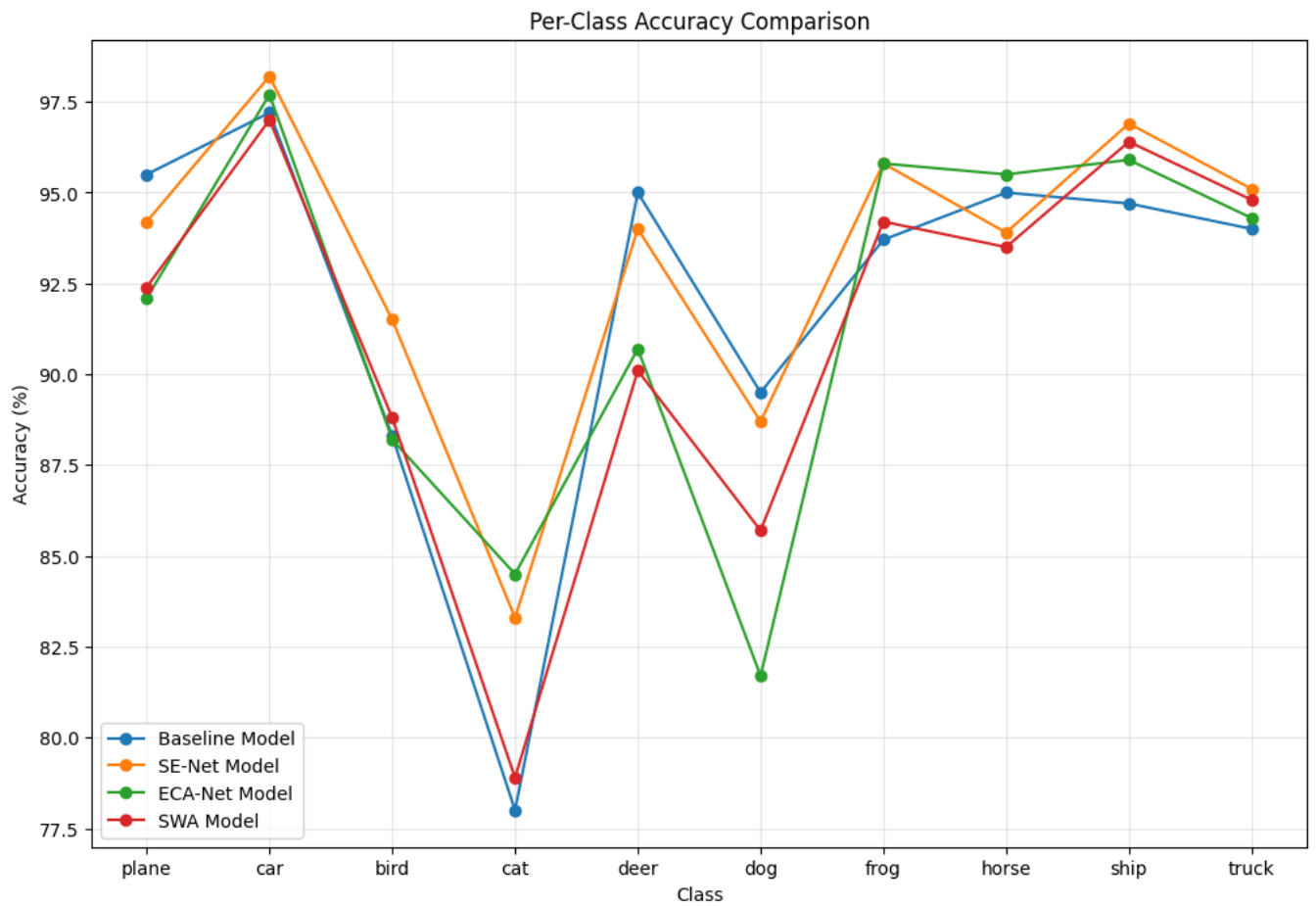
Model Training Process Comparison



Best Model: SE-Net Model with accuracy 93.16%

SE-Net Model Confusion Matrix





=== All Experiments Completed ===

Baseline Model Final Accuracy: 92.09%

SE-Net Model Final Accuracy: 93.16%

ECA-Net Model Final Accuracy: 91.64%

SWA Model Final Accuracy: 91.18%

Best accuracy improvement: 1.07%p