**Result:**

**Execution Time Analysis**

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Execution Time (Baseline, ms)** | **Execution Time (HAF-CNN, ms)** | **Speed Up** |
| **Baseline** | **HAF CNN** | **%** |
| **ResNet-50** | **320** | **245** | **23.4375** |
| **EfficientNet-B0** | **610** | **460** | **24.590164** |
| **MobileNetV2** | **120** | **95** | **20.833333** |
| **VGG-16** | **400** | **310** | **22.5** |
| **AlexNet** | **250** | **195** | **22** |
| **DenseNet-121** | **460** | **360** | **21.73913** |
| **Inception-v3** | **520** | **405** | **22.115385** |

Data Transfer Time Analysis between Host and Device

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Model | H2D Transfer Time (ms) | H2D Transfer Time HAF-CNN (ms) | H2D Transfer Time Speed(%) | D2H Transfer Time (ms) | D2H Transfer Time HAF-CNN (ms) | D2H Transfer Time Speed(%) |
|  | Baseline | HAF CNN | Baseline | HAF CNN |
| ResNet-50 | 45 | 32 | 28.888889 | 41 | 30 | 26.82927 |
| EfficientNet-B0 | 52 | 40 | 23.076923 | 49 | 35 | 28.57143 |
| MobileNetV2 | 20 | 14 | 30 | 18 | 13 | 27.77778 |
| VGG-16 | 48 | 36 | 25 | 44 | 33 | 25 |
| AlexNet | 22 | 16 | 27.272727 | 20 | 15 | 25 |
| DenseNet-121 | 50 | 38 | 24 | 47 | 34 | 27.65957 |
| Inception-v3 | 55 | 40 | 27.272727 | 51 | 37 | 27.45098 |

GPU Utilization Analysis

|  |  |  |  |
| --- | --- | --- | --- |
| Model | GPU Utilization (Baseline, %) | GPU Utilization (HAF-CNN, %) | GPU Utilization Improvement (%) |
|  | Baseline | HAF CNN |
| ResNet-50 | 68 | 86 | 18 |
| EfficientNet-B0 | 65 | 80 | 15 |
| MobileNetV2 | 60 | 82 | 22 |
| VGG-16 | 70 | 85 | 15 |
| AlexNet | 58 | 76 | 18 |
| DenseNet-121 | 68 | 84 | 16 |
| Inception-v3 | 66 | 81 | 15 |

|  |  |  |
| --- | --- | --- |
|  | Peak Memory Usage Analysis |  |

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Peak Memory (Baseline, GB) | Peak Memory (HAF-CNN, GB) | Peak Memory Usage Improvement (GB) |
|  | Baseline | HAF CNN |
| ResNet-50 | 4.8 | 4.1 | 0.7 |
| EfficientNet-B0 | 5.4 | 4.2 | 1.2 |
| MobileNetV2 | 2 | 1.6 | 0.4 |
| VGG-16 | 5 | 4.1 | 0.9 |
| AlexNet | 2.3 | 1.8 | 0.5 |
| DenseNet-121 | 4.5 | 3.6 | 0.9 |
| Inception-v3 | 4.8 | 3.9 | 0.9 |

Kernel Launch Count Analysis

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Kernel Launch Count (Baseline) | Kernel Launch Count (HAF-CNN) | Kernel Launch Counts Difference(%) |
|  | Baseline | HAF CNN |
| ResNet-50 | 320 | 138 | 56.875 |
| EfficientNet-B0 | 500 | 260 | 48 |
| MobileNetV2 | 180 | 95 | 47.222222 |
| VGG-16 | 420 | 190 | 54.761905 |
| AlexNet | 220 | 120 | 45.454545 |
| DenseNet-121 | 480 | 240 | 50 |
| Inception-v3 | 520 | 265 | 49.038462 |

Power Consumption Analysis

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Power Consumption (Baseline, W) | Power Consumption (HAF-CNN, W) | Energy Efficiency(%) |
|  | Baseline | HAF CNN |
| ResNet-50 | 280 | 250 | 10.714286 |
| EfficientNet-B0 | 300 | 265 | 11.666667 |
| MobileNetV2 | 230 | 210 | 8.6956522 |
| VGG-16 | 310 | 275 | 11.290323 |
| AlexNet | 240 | 220 | 8.3333333 |
| DenseNet-121 | 295 | 260 | 11.864407 |
| Inception-v3 | 320 | 285 | 10.93 |

Accuracy Validation

|  |  |
| --- | --- |
| Model | Accuracy Difference Between Baseline & HAF CNN (%) |
|
| ResNet-50 | 0.2 |
| EfficientNet-B0 | 0.2 |
| MobileNetV2 | 0.1 |
| VGG-16 | 0.3 |
| AlexNet | 0.1 |
| DenseNet-121 | 0.2 |
| Inception-v3 | 0.2 |