

Classify up to 1.21x the Frames Per Second for ResNet50 Workloads by Choosing AWS M6i Instances with 3rd Gen Intel[®] Xeon[®] Scalable Processors

M6i Instances Accelerated Image Classification Over M5n Instances with Previous-Generation Processors

Running deep learning workloads in the cloud can give organizations the power to make sense of their data while also getting the flexibility and easy scalability the cloud provides. For those using the ResNet50 framework to classify images, Principled Technologies compared the performance of two instance types:

- AWS M6i instances with 3rd Gen Intel[®] Xeon[®] Scalable processors
- AWS M5n instances with 2nd Gen Intel Xeon Scalable processors

Using int8 precision and a batch size of 128 on the ResNet50 benchmark, results showed that for both small and large instance sizes, AWS M6i instances offered up to 1.21x the frames per second rate for image classification. This means that if your organization seeks to run these types of image classification workloads in the cloud, AWS M6i instances can cut through the data faster to deliver insights sooner.

Making Sense of Data Faster on Large Instances

First, to show how workloads with higher throughput demands might compare, Principled Technologies tested with large, 96 vCPU instances. The updated AWS M6i instances with 3rd Gen Intel[®] Xeon[®] Scalable processors offered a significant performance advantage over the M5n instances with previous-generation processors, handling 1.21x the frames per second on the ResNet50 benchmark (see Figure 1).

Large instance ResNet50 workload comparison

Frames per second (normalized) | Higher is better

Precision: int8
Batch size: 128

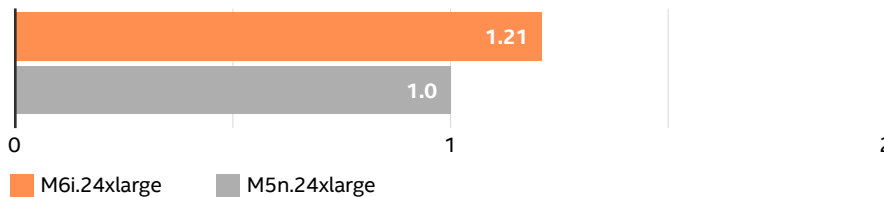


Figure 1. The relative rate of frames per second for large M6i and M5n instances (96 vCPUs) completing image classification using the ResNet50 benchmark. Higher numbers are better.

ResNet50

Classify 1.21x the frames per second on 96 vCPU AWS M6i instances with 3rd Gen Intel Xeon Scalable processors
vs. M5n instances

Classify 1.19x the frames per second on 16 vCPU AWS M6i instances with 3rd Gen Intel Xeon Scalable processors
vs. M5n instances

Making Sense of Data Faster on Small Instances

As Figure 2 shows, small instances with 16 vCPUs offered similar performance deltas as the larger instances with 96 vCPUs. Test results show that AWS M6i instances with 3rd Gen Intel® Xeon® processors classified 1.19x the frames per second of M5n instances with previous-generation processors.

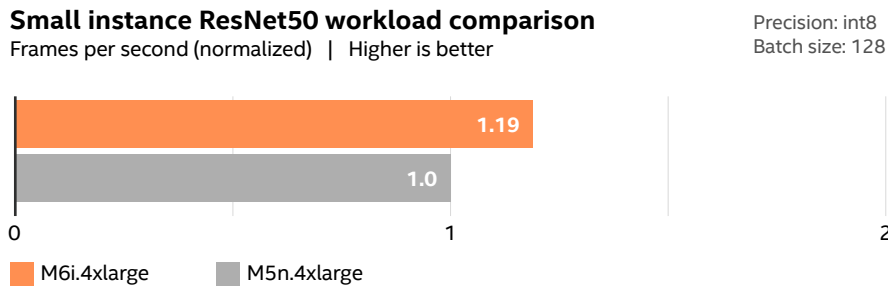


Figure 2. The relative rate of frames per second for small M6i and M5n instances (16 vCPUs) completing image classification using the ResNet50 benchmark. Higher numbers are better.

Conclusion

Using deep learning frameworks for image classification can offer new ways to diagnose medical conditions or assess damages to city infrastructure after a hurricane. No matter your particular use case for running ResNet50 to classify images, getting through more frames per second can lead to insights. Tests show that opting for AWS M6i instances with 3rd Gen Intel Xeon Scalable processors over M5n instances with previous-generation processors can improve the frames per second rate for image classification by as much as 1.21x.

Analyzing images at a faster rate has more implications than just getting answers more quickly. If the cloud instance type you choose can process more images, faster, it can reduce the number of instances you need to purchase—which is also a win for your bottom line.

Learn More

To begin running your image classification workloads on AWS M6i instances, visit <https://aws.amazon.com/ec2/instance-types/m6i/>.

For more details about these third-party test results, visit <https://facts.pt/DNLGDAd>.



Performance varies by use, configuration and other factors. Learn more at www.intel.com/PerformanceIndex.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure. Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

Printed in USA 0322/JO/PT/PDF US001

Please Recycle