



On-Device Artificial Intelligence
Eunsoo Shim, SAIT, Samsung Electronics

The world has entered the era of AI where AI is expected to have high impact on numerous sectors of our society.

It is known that big data, large computing power (in particular, provided by GPUs), and deep learning are three

main ingredients enabling the current rapid progress of AI.

Most of the AI applications have been deployed in the cloud due to its demand for large computing power.

A new trend is that there is strong demand for deploying AI applications in embedded systems without relying on the cloud.

We call such AI technologies on-device AI.

In this talk, we review hardware technologies and algorithms for on-device AI.