Motivation

* Speicherbedarf : CNN haben Millionen Parameter, Geräte eingeschränkter Speicherplatz.
* Große Anzahl von Speicherzugriffen und arithmetischen Operationen->Strom Verbrauch

Pruning

* Sparsamkeit im CNN: Reduktion der nicht-nulle Parameter.
* CNNs haben überflüssige Parameter.
* Kleiner Verlust der Genauigkeit bei großen CNNs.
* Großer Verlust der Genauigkeit bei kleinen CNNs.