**GPPs** FPGA Accelerator Forward Propagation Distillation Reference Fixed-point Model(teacher) Teacher Network Softmax  $\mathbf{z}^{\mathrm{T}}$ p<sup>T</sup>(Soft-Target) **CNN** Configuration 000 ► SoftmaxWithLoss Soft-Loss Student Network SoftmaxWithLoss Input Image data 000 Hard-Loss Initial Fixed-point Model(student) Student-Loss **Backward Propagation** Hard-Target Image Label Student Network Update Weights 000