

Knowledge distillation aims at transferring knowledge acquired in one model (a teacher) to another model (a student) that is typically smaller. Previous approach