Codebook初始化：

C1: 16x64的初始codebook

C2: 16x128的自注意力处理后的codebook

重参数化采样：

C3: (batchsize, 16, 128)，从C2采样得到

生成器G1：

输入：C3

输出：F1 (batchsize, 128, 128)

功能：分16个区域生成128x128的图片

真实数据：

P1: (batchsize, 128, 128)

抽取器E1：

输入：F1或P1

输出：RF1或RP1 (batchsize, 16, 128)

功能：将输入划分为16个区域，每个区域用CNN还原为128维向量

判别器C1：

输入：RF1或RP1的16个向量

功能：判断每个向量是真实还是伪造

Codebook替换：

使用C2替换RF1和RP1中的向量，得到TRF和TRP

计算替换损失L\_TRP

重构过程：

G1(TRP) -> GTRP (batchsize, 128, 128)

计算重构损失L\_R\_GTRP：GTRP与P1的差异

G1(TRF) -> GTRF

二次生成和判别：

E1(GTRF) -> RGTRF

E1(GTRP) -> RGTRP

E1(GTRF) -> RGTRF

C1判别RGTRF和RGTRP

C2替换（RGTRP）->TRGTRP

C2替换（RGTRF）->TRGTRF

计算替换损失L\_TRGRP

G1(TRGTRP)->GTRGTRP

计算重构损失L\_ GTRGTRP

G1(TRGTRF) ->GTRGTRF

训练目标：

判别器C1：

将RF1, RGTRF判断为伪造

将RP1, RGTRP判断为真实

生成器G1：

欺骗C1，使RF1, RGTRF被判断为真实

最小化重构损失L\_R\_GTRP，L\_ GTRGTRP

整体：

最小化codebook替换损失L\_TRP，L\_TRGRP