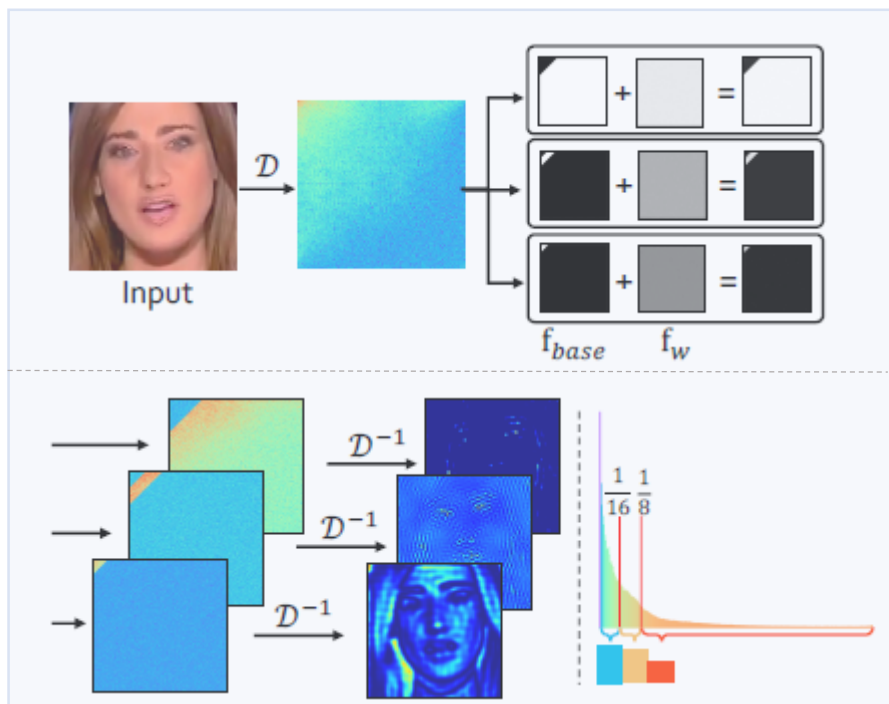
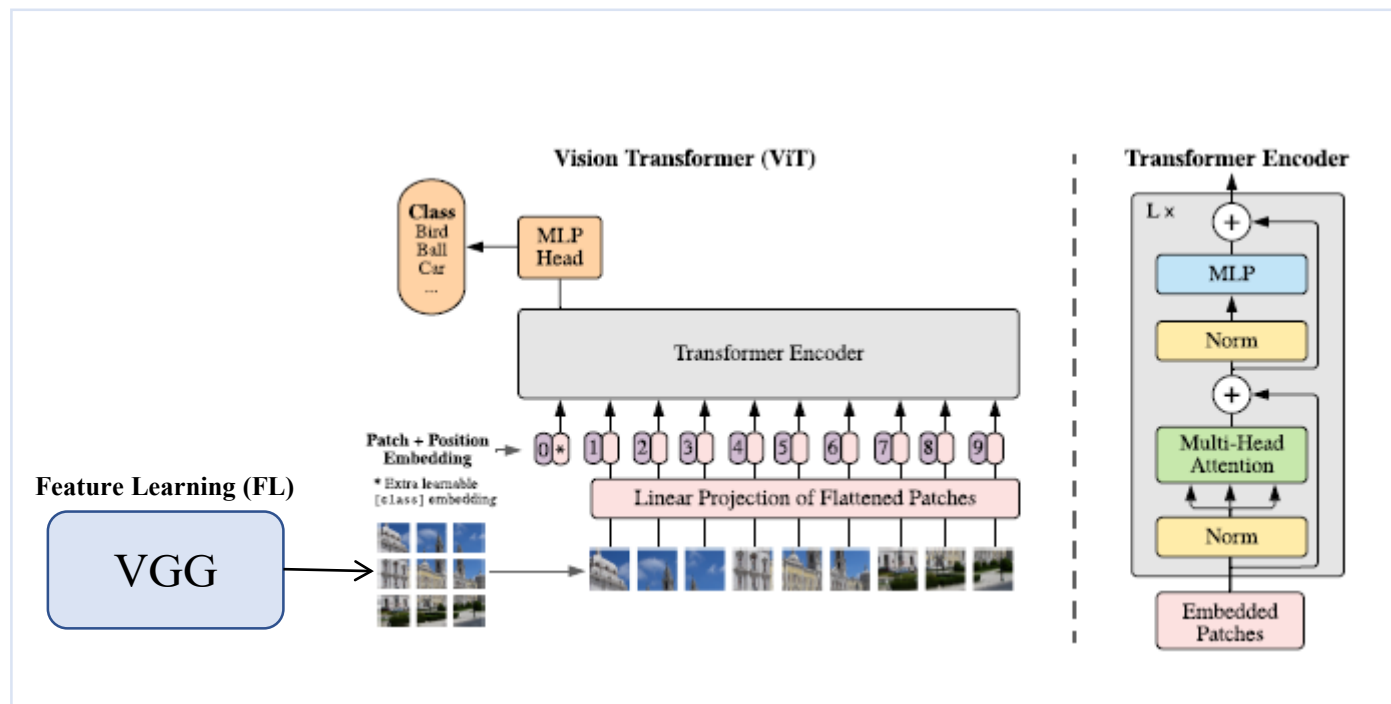


● 频域 ---> CViT



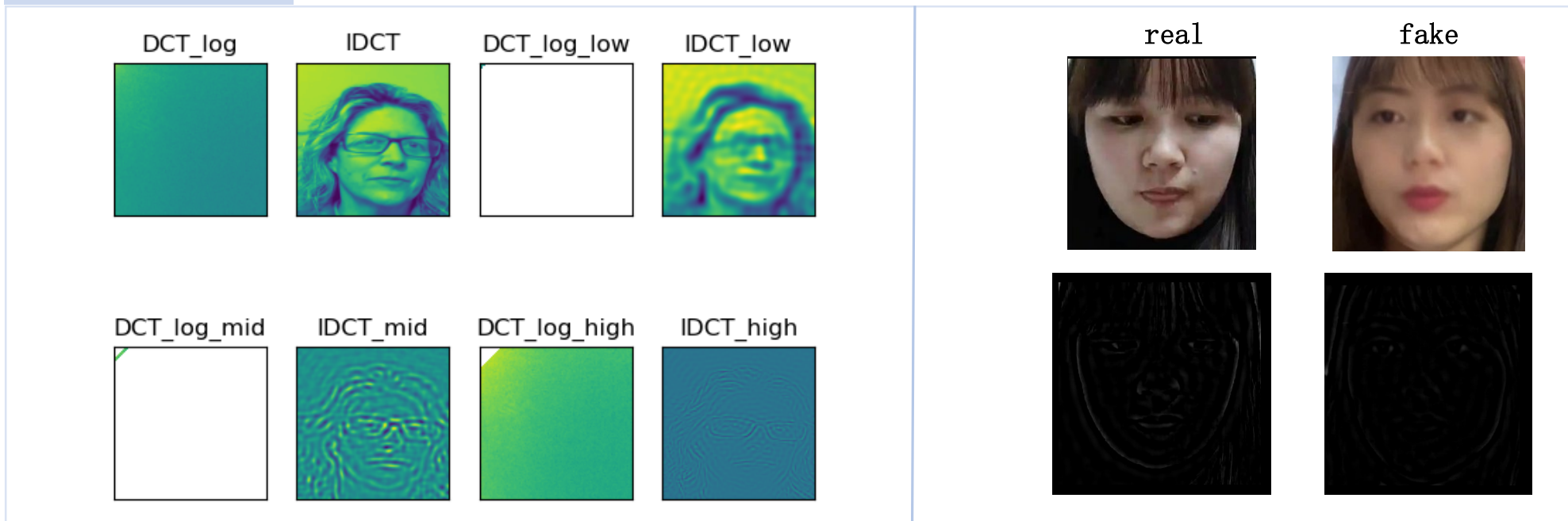
DCT离散余弦变换-->频域：分别保留低频、中频、高频信息
DCT逆变换-->三种频率域下恢复原图像
低频揭示全局特征，中高频揭示细节伪影



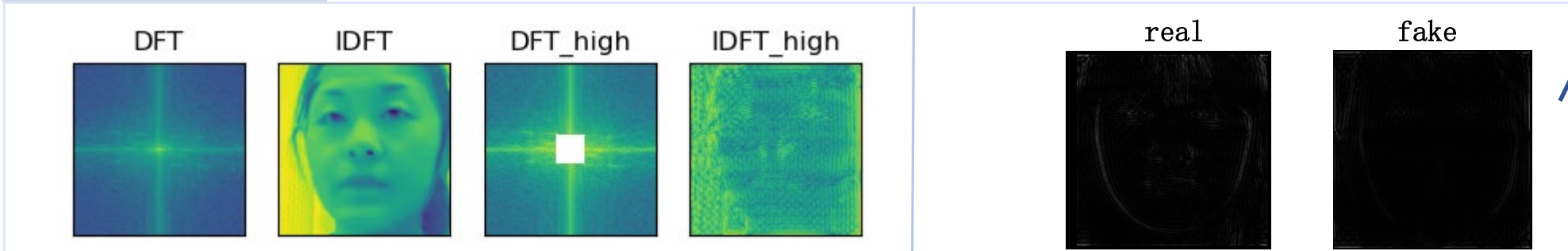
FL (没有全连接层的VGG架构，不进行分类) 提取可学习的特征
ViT将学习到的特征作为输入，并用注意机制进行分类

1. Qian, Y., Yin, G., Sheng, L., Chen, Z., & Shao, J. (2020, August). **Thinking in frequency: Face forgery detection by mining frequency-aware clues**. In European Conference on Computer Vision (pp. 86-103).
2. Wodajo, D., & Atnafu, S. (2021). **Deepfake Video Detection Using Convolutional Vision Transformer**. arXiv preprint arXiv:2102.11126.

DCT离散余弦变换



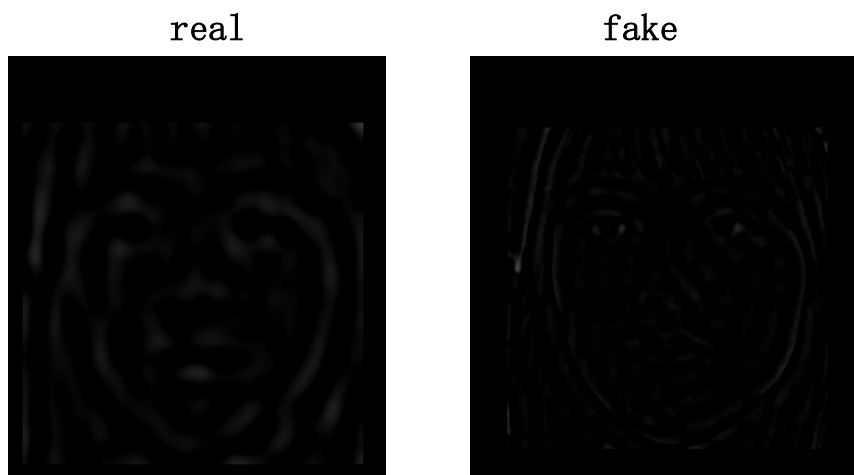
DFT离散傅里叶变换



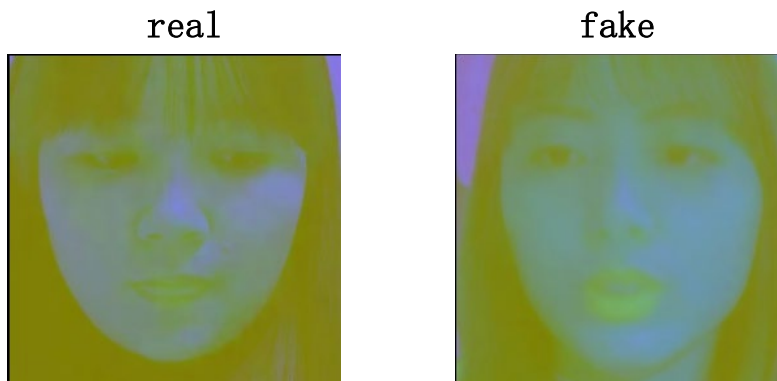
只保留高频信息时，真、假图片有差异

- 频域 + 颜色空间

- DFT频域空间（高频信息）



- YCbCr/HSV颜色空间（亮度与色度分离）



cross-domain

FMFCC比赛

1. 训练数据少
每个视频取关键帧、10帧
2. 网络架构
CNN + transform 不适用
3. 数据增强
原始图片增强，再转换