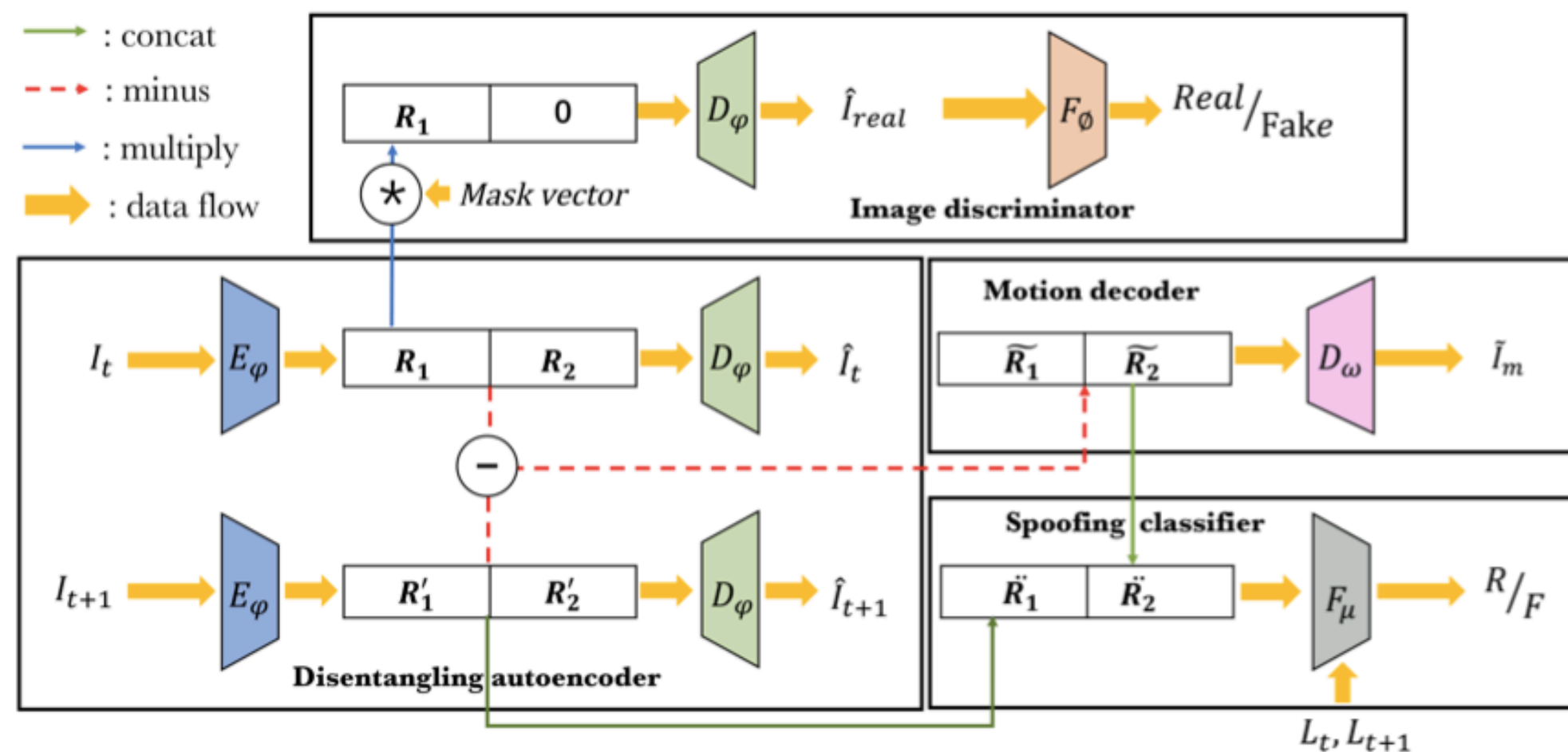


Disentangled Representation based Face Anti-Spoofing

Zhao Liu*, Zunlei Feng, Yong Li, Zeyu Zou, Rong Zhang, Mingli Song, Jianping Shen

Face anti-spoofing is an important problem for both academic research and industrial face recognition systems. Most of the existing face anti-spoofing methods take it as a classification task on individual static images, where motion pattern differences in consecutive real or fake face sequences are ignored. In this work, we propose a novel method to identify spoofing patterns using motion information. Different from previous methods, the proposed method makes the real or fake decision on the disentangled feature level, based on the observation that motion and spoofing pattern features could be disentangled from original image frames. We design a representation disentangling framework for this task, which is able to reconstruct both real and fake face sequences from the input. Meanwhile, the disentangled representations could be used to classify whether the input faces are real or fake. We perform several experiments on Casia-FASD and ReplayAttack datasets. The proposed method achieves SOTA results compared with existing face anti-spoofing methods.



Algorithm 1 Training and inference process

Training

Input: Face video sequences $\{V_A = v_1, v_2, \dots, v_N\}$, labels $\{L_A = l_1, l_2, \dots, l_N\}$, number of epoches k

- 1: **while** i is less than k **do**
- 2: Choose image pairs $(I_t^1, I_{t+1}^1), \dots, (I_t^N, I_{t+1}^N)$, with labels l_1, \dots, l_N
- 3: **for** each image pair (I_t^k, I_{t+1}^k) and its label l_k **do**
- 4: Encode (I_t^k, I_{t+1}^k) into (R, R')
- 5: Decode (R, R') into images $(\hat{I}_t^k, \hat{I}_{t+1}^k)$
- 6: Reconstruct image pairs $(\hat{I}_{real}, \hat{I}_{real})$
- 7: Disentangle motion pattern map I_m
- 8: Update the parameters of functional networks based on the computed losses L_g
- 9: Combine the disentangled motion pattern and spoofing pattern
- 10: Classify the spoofing type
- 11: Update the parameters of spoofing classifier based on classification result and input label l_k

Output: Pretrained functional networks

Inference

Input: Test face video sequence v_T

- 1: Choose image pair (I_t^T, I_{t+1}^T) from v_T
- 2: Disentangle the motion pattern and spoofing pattern
- 3: Decide the spoofing type by spoofing classifier

Output: Whether v_T is a real or spoofing face sequence

Visualization movement on Casia Dataset

