DETECTING FACE FORGERY BY LEARNING THE RELATIONSHIPS BETWEEN BASIC FACIAL ACTION UNITS

Đinh Hoàng Thùy Linh

Trường Đại học Công nghệ Thông tin - Đại học Quốc gia TP.HCM

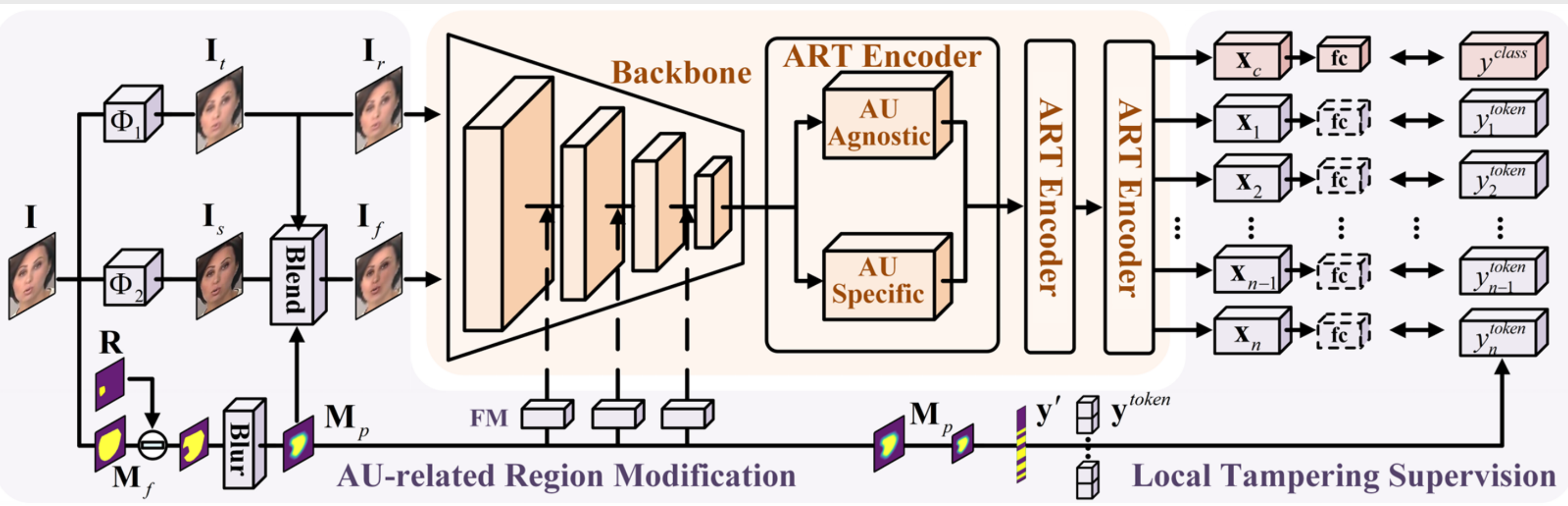
What ?

ART Module: Learns AU relationships via:
• AU-specific branch (attention-based AU interactions)
• AU-agnostic branch (Vision Transformer for patch relations)
TAP Module: Boosts manipulation sensitivity by:
• Random AU region masking
• Multi-level fake sample generation
• Location-aware tampered area prediction

Why ?

- GAN-based face forgery techniques are increasingly realistic, challenging detection.
- Current methods work well on known manipulations but fail on unseen types.
- We exploit disrupted facial Action Unit (AU) relationships—key natural muscle movement patterns—as detection clues.
- Our framework combines data augmentation and auxiliary tasks to boost accuracy and generalization.

Overview



Description

1. Summary

The proposed Action Units Relation Learning framework consists of two main components: ART (Action Units Relation Transformer) and TAP (Tampered AU Prediction).

- ART learns the relationships between AU-related regions to enhance forgery detection.
- TAP tampers with AU-related regions and provides Local Tampering Supervision to improve the model's generalization ability.

2. Action Units Relation Transformer

Action Units Relation Transformer (ART) consists of three stacked encoders, each comprising two branches:

- The AU-specific branch focuses on regions associated with individual Action Units (AUs), using attention mechanisms to model their relationships.
- The AU-agnostic branch employs a Transformer to capture global dependencies among image patches, including both AU-related and other facial features.
- The outputs from both branches—representing fine-grained and global relational features—are fused via element-wise addition and convolutional layers to produce the final representation for forgery detection.

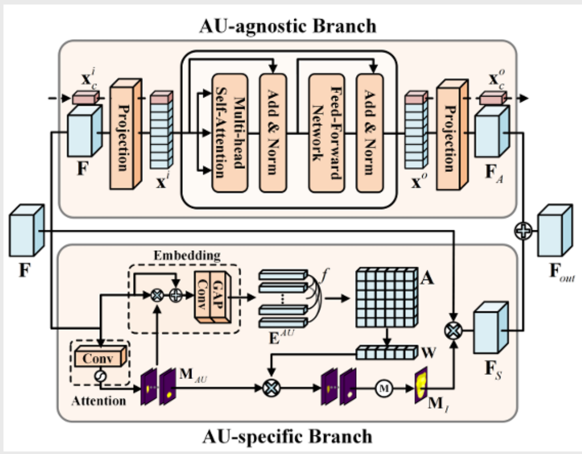


Figure 3. The details of our proposed ART Encoder..

3.Tampered AU Prediction

TAP is an auxiliary task designed to enhance the detection of localized facial tampering, specifically targeting Action Units (AU)-related regions. TAP consists of two main components:

3.1. TheAU-related Region Modification (ARM):

- Image-level Tampering: Generates fake images by applying Gaussian blur to randomly selected AU regions based on facial landmark-derived masks.
- Feature-level Mixing: Blends statistical features from real images into fake images at a randomly selected layer, applied only to tampered regions to create more challenging synthetic samples.

3.2. Local Tampering Supervision (LS):

- Instead of relying solely on a global [CLS] token for forgery detection, LS assigns individual supervision labels to each image patch token, reflecting tampering likelihood.
- The model learns to predict manipulated regions through these localized tokens.
- During inference, predictions are made by combining both the global [CLS] token and local patch tokens.

NII

Đinh Hoàng Thùy Linh – Trường Đại học Công Nghệ Thông Tin

TEL : 0989531477 Email : linhdht.19@grad.uit.edu.vn