Novelty Assessment Report

Paper: RelayFormer: A Unified Local-Global Attention Framework for Scalable Image and Video Manipulation Localization

PDF URL: https://openreview.net/pdf?id=e61YQdLIam

Venue: ICLR 2026 Conference Submission

Year: 2026

Report Generated: 2026-01-04

Abstract

Visual manipulation localization (VML) aims to identify tampered regions in images and videos, a task that has become increasingly challenging with the rise of advanced editing tools. Existing methods face two main issues: resolution diversity, where resizing or padding distorts forensic traces and reduces efficiency, and the modality gap, as images and videos often require separate models. To address these challenges, we propose RelayFormer, a unified framework that adapts to varying resolutions and modalities. RelayFormer partitions inputs into fixed-size sub-images and introduces Global-Local Relay (GLR) tokens, which propagate structured context through a global-local relay attention (GLRA) mechanism. This enables efficient exchange of global cues, such as semantic or temporal consistency, while preserving fine-grained manipulation artifacts. Unlike prior methods that rely on uniform resizing or sparse attention, RelayFormer naturally scales to arbitrary resolutions and video sequences without excessive overhead. Experiments across diverse benchmarks demonstrate that RelayFormer achieves state-of-the-art performance with notable efficiency, combining resolution adaptivity without interpolation or excessive padding, unified modeling for both images and videos, and a strong balance between accuracy and computational cost.

Disclaimer

This report is **AI-GENERATED** using Large Language Models and WisPaper (a scholar search engine). It analyzes academic papers' tasks and contributions against retrieved prior work. While this system identifies **POTENTIAL** overlaps and novel directions, **ITS COVERAGE IS NOT EXHAUSTIVE AND JUDGMENTS ARE APPROXIMATE**. These results are intended to assist human reviewers and **SHOULD NOT** be relied upon as a definitive verdict on novelty.

Note that some papers exist in multiple, slightly different versions (e.g., with different titles or URLs). The system may retrieve several versions of the same underlying work. The current automated pipeline does not reliably align or distinguish these cases, so human reviewers will need to disambiguate them manually.

If you have any questions, please contact: mingzhang23@m.fudan.edu.cn

Core Task Landscape

This paper addresses: **Visual Manipulation Localization in Images and Videos**

A total of **50 papers** were analyzed and organized into a taxonomy with **12 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Deepfake and Face Forgery Detection**
- **Image Tampering Detection and Localization**
- **Video Tampering Detection and Localization**
- **Watermarking-Based Authentication and Tamper Detection**
- **Unified and Scalable Manipulation Localization Frameworks**
- **Surveys, Reviews, and Application-Specific Studies**

Complete Taxonomy Tree

- Visual Manipulation Localization in Images and Videos Survey Taxonomy
- Deepfake and Face Forgery Detection
  - Spatial-Temporal Deepfake Detection (3 papers)
  - [4] Deepfake Detection Using Convolutional Neural Networks and LSTM Modelling (Sheshang Degadwala, 2025) View paper
  - [17] Forgery Detection Scheme of Deep Video Frame-rate Up-conversion Based on Dual-stream Multi-scale Spatial-temporal Representation (Qing Gu, 2022) View paper
  - [22] SRTNet: a spatial and residual based two-stream neural network for deepfakes detection (Dengyong Zhang, 2022) View paper
  - Spatial or Noise-Based Deepfake Detection (3 papers)
  - [3] Face Forgery Detection via Multi-Feature Fusion and Local Enhancement (Dengyong Zhang, 2024) View paper
  - [16] Improving Generalization in Facial Manipulation Detection Using Image Noise Residuals and Temporal Features (Mehdi Atamna, 2023) View paper
  - [32] Detecting DeepFake, FaceSwap and Face2Face facial forgeries using frequency CNN (Aditi Kohli, 2021) View paper
  - General Deepfake Detection Frameworks (3 papers)
  - [12] An efficient deepfake video detection using robust deep learning (Abdul Qadir, 2024) View paper
  - [25] An efficient deepfake detection using robust deep learning approch (Abdul Qadir, 2023) View paper
  - [48] Deep-Fake Visual Detection using AI (Dhage, 2025) View paper
- Image Tampering Detection and Localization
  - Deep Learning-Based Image Tampering Detection (6 papers)
  - [2] Detection of image tampering using multiscale fusion and anomalousness assessment (Yichen Wang, 2024) View paper
  - [18] An interpretable image tampering detection approach based on cooperative game (Wei Lu, 2022) View paper
  - [26] Frequency-spatial feature integration with boundary-aware learning for image tampering detection (D Dagar, 2025) View paper
  - [27] Image Tampering Detection (Sandeep Shinde, 2025) View paper
  - [29] Image Manipulation Localization Using Attentional Cross-Domain CNN Features (Shuaibo Li, 2021) View paper
  - [30] Tampering Detection and Segmentation Model for Multimedia Forensic (Manjunatha. S, 2023) View paper
  - Traditional or Keypoint-Based Image Tampering Detection (2 papers)
  - [24] A deep neural network with hybrid spotted hyena optimizer and grasshopper optimization algorithm for copy move forgery detection (R. Gupta, 2022) View paper

- [47] Image integrity and tampering detection: A hybrid approach to copy-paste forgery detection using ORB-SSD and CNN (Priti Badar, 2025) View paper
- Domain-Specific Image Tampering Detection (3 papers)
- [10] A novel blind tamper detection and localization scheme for multiple faces in digital images (Rasha Thabit, 2023) View paper
- [14] An exhaustive review of authentication, tamper detection with localization and recovery techniques for medical images (B. Madhushree, 2023) View paper
- [23] VizDefender: Unmasking Visualization Tampering through Proactive Localization and Intent Inference (Sicheng Song, 2025) View paper

- Video Tampering Detection and Localization
  - Passive Video Tampering Detection (3 papers)
  - [20] Video tamper detection based on multi-scale mutual information (Wei Wei, 2019) View paper
  - [35] Local tampering detection in video sequences (Paolo Bestagini, 2013) View paper
  - [42] Malicious inter-frame video tampering detection in MPEG videos using time and spatial domain analysis of quantization effects (Javad Abbasi Aghamaleki, 2017) View paper
- Watermarking-Based Authentication and Tamper Detection
  - Image Watermarking for Tamper Detection (8 papers)
  - [1] A study on content tampering in multimedia watermarking (Aditya Kumar Sahu, 2023) View paper
  - [5] Image tamper detection and self-recovery using multiple median watermarking (Vishal Rajput, 2020) View paper
  - [6] Chaotic watermarking for tamper detection: Enhancing robustness and security in digital multimedia (H Kaur, 2024) View paper
  - [28] A novel chaos-based fragile watermarking for image tampering detection and self-recovery (Xiaojun Tong, 2013) View paper
  - [31] Image watermarking for tamper detection (Jessica Fridrich, 1998) View paper
  - [36] A hybrid-Sudoku based fragile watermarking scheme for image tampering detection (Guo-Dong Su, 2021) View paper
  - [37] A dual-embedded tamper detection framework based on block truncation coding for intelligent multimedia systems (Mianjie Li, 2023) View paper
  - [40] Large scale image tamper detection and restoration (Debojit Sarkar, 2020) View paper
  - Video Watermarking for Tamper Detection (5 papers)
  - [21] Blind Semi-fragile Hybrid Domain-Based Dual Watermarking System for Video Authentication and Tampering Localization (Amal Hammami, 2023) View paper
  - [33] Multimedia tamper detection and localization using digital watermarking (Aditya Arsh, 2024) View paper
  - [38] Tampering detection in compressed digital video using watermarking (M. Fallahpour, 2014) View paper
  - [39] Tampering Detection of Audio-Visual Content Using Encrypted Watermarks (Ronaldo Rigoni, 2014) View paper
  - [49] Video-tampering detection and content reconstruction via self-embedding (Vahideh Amanipour, 2017) View paper
- Unified and Scalable Manipulation Localization Frameworks ★ (1 papers)
  - [0] RelayFormer: A Unified Local-Global Attention Framework for Scalable Image and Video Manipulation Localization (Anon et al., 2026) View paper
- Surveys, Reviews, and Application-Specific Studies
  - General Surveys and Reviews (9 papers)
  - [7] Towards Analysis Detection of Deepfake Video via Deep Learning Models: A Review (Shahad Altamimi, 2024) View paper
  - [8] Fighting fake visual media: a study of current and emerging methods for detecting image and video tampering (Mahejabi Khan, 2024) View paper
  - [9] Unravelling digital forgeries: A systematic survey on image manipulation detection and localization (VijayaKumar Kadha, 2025) View paper
  - [11] A detailed analysis of image and video forgery detection techniques (Shobhit Tyagi, 2023) View paper
  - [13] Visualizing the truth: a survey of multimedia forensic analysis (Anjali Diwan, 2024) View paper
  - [15] An overview of video tampering detection techniques: state-of-the-art and future directions (Saroj Kumar Pandey, 2023) View paper
  - [19] A Comprehensive Survey on Detection of Fake Multimedia Content (Ayush S Acharya, 2025) View paper
  - [43] Digital video tampering detection: An overview of passive techniques (K. Sitara, 2016) View paper
  - [46] Securing Digital Media Integrity: A Survey of Watermarking and Manipulation Detection for Image Authentication (Liu Xinyun, 2025) View paper
  - Application-Specific Studies (5 papers)
  - [34] Robust image alignment for tampering detection (Sebastiano Battiato, 2012) View paper
  - [41] News authentication and tampered images: evaluating the photo-truth impact through image verification algorithms (Anastasia Katsaounidou, 2020) View paper
  - [44] VisGuard: Securing Visualization Dissemination through Tamper-Resistant Data Retrieval (Chen, 2025) View paper
  - [45] TAMPAR: Visual Tampering Detection for Parcel Logistics in Postal Supply Chains (Alexander Naumann, 2024) View paper
  - [50] Camera Tampering Detection using Generative Reference Model and Deep Learned Features. (Pranav Mantini, 2019) View paper

Narrative

Core task: visual manipulation localization in images and videos. The field has evolved into several distinct branches that reflect both the nature of the forgery and the detection strategy. Deepfake and Face Forgery Detection focuses on synthetic face generation and identity swaps, often leveraging temporal cues and multi-feature fusion as seen in Face Forgery Multi-Feature[3] and Deepfake CNN LSTM[4]. Image Tampering Detection and Localization addresses copy-move, splicing, and inpainting operations through methods that exploit noise residuals, frequency artifacts, or learned segmentation models such as Tampering Segmentation Model[30] and Multiscale Fusion Detection[2]. Video Tampering Detection and Localization extends these ideas to temporal domains, examining frame insertion, deletion, and inter-frame inconsistencies with approaches like Noise Residuals Temporal[16] and Frame-rate Forgery Detection[17]. Watermarking-Based Authentication embeds fragile or semi-fragile signals into content for tamper localization, exemplified by Multiple Median Watermarking[5] and Chaotic Watermarking[6]. Unified and Scalable Manipulation Localization Frameworks aim to handle diverse forgery types within a single architecture, while Surveys, Reviews, and Application-Specific Studies provide overviews and domain-specific analyses, including Deepfake Analysis Review[7] and Fighting Fake Media[8].

Recent work has increasingly emphasized cross-domain generalization and the integration of spatial and frequency features to improve robustness against unseen manipulations. A handful of studies explore hybrid strategies that combine passive forensic cues with active watermarking signals, as in Content Tampering Watermarking[1] and Dual-Embedded Framework[37]. RelayFormer[0] sits within the Unified and Scalable Manipulation Localization Frameworks branch, proposing a transformer-based architecture designed to localize manipulations across multiple forgery types without retraining for each specific attack. This contrasts with more specialized detectors like Efficient Deepfake Detection[12], which targets face forgeries exclusively, and with watermarking methods such as Chaotic Watermarking[6], which require embedding at capture time. By aiming for a single model that generalizes across image and video tampering scenarios, RelayFormer[0] addresses a key open question: how to scale forensic systems to the growing diversity of generative and editing tools without sacrificing localization precision.

Related Works in Same Category

No sibling papers and no sibling subtopics were found under the same parent taxonomy node; the paper appears structurally isolated in the taxonomy.

Contributions Analysis

Overall novelty summary. RelayFormer proposes a unified framework for visual manipulation localization that adapts to varying resolutions and modalities through fixed-size sub-image partitioning and Global-Local Relay tokens. The paper resides in the 'Unified and Scalable Manipulation Localization Frameworks' leaf, which contains only this single work among the 50 papers surveyed. This isolation suggests the research direction—addressing resolution diversity and modality gaps simultaneously within one architecture—is relatively unexplored in the current taxonomy, positioning the work in a sparse rather than crowded area of the field.

The taxonomy reveals that neighboring leaves focus on specialized detection strategies: 'Deep Learning-Based Image Tampering Detection' contains six papers emphasizing spatial or frequency features for still images, 'Passive Video Tampering Detection' includes three papers analyzing compression artifacts and motion residuals, and 'Spatial-Temporal Deepfake Detection' groups three papers combining CNN-LSTM architectures for face forgeries. RelayFormer diverges by targeting cross-modality generalization rather than optimizing for a single forgery type or medium, bridging gaps that prior work addresses through separate models or fixed-resolution preprocessing.

Among 25 candidates examined, the unified framework contribution (10 candidates, 0 refutable) and the GLR token mechanism (5 candidates, 0 refutable) show no clear prior overlap within the limited search scope. The query-based mask decoder (10 candidates, 1 refutable) encounters one candidate suggesting overlapping prior work, indicating this component may have precedent in segmentation or detection literature. The statistics reflect a modest search scale—top-K semantic matches plus citation expansion—so the absence of refutation for two contributions does not guarantee exhaustive novelty but suggests limited direct precedent among closely related papers.

Based on the limited search scope, RelayFormer appears to occupy a relatively novel position by unifying resolution adaptivity and modality handling in a single framework, though the query-based decoder component may have more substantial prior work. The taxonomy structure and contribution-level statistics together suggest the core relay-token mechanism and unified architecture are less explored, while acknowledging that the 25-candidate search cannot rule out relevant work outside the examined set.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: RelayFormer unified framework for resolution-adaptive manipulation localization

Description: The authors introduce RelayFormer, a framework that processes images and videos of arbitrary resolutions without interpolation or padding by partitioning inputs into fixed-size sub-images. This unified architecture handles both image and video manipulation localization tasks within a single model.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. MSER-Net: Multi-stage edge refinement network for deepfake detection

URL: View paper

Brief Assessment

MSER-Net[71] focuses on deepfake detection using multi-scale edge refinement with Sobel operators, which is a different technical approach from RelayFormer's resolution-adaptive framework using sub-image partitioning and relay tokens for manipulation localization.

2. Refining localized attention features with multi-scale relationships for enhanced deepfake detection in spatial-frequency domain

URL: View paper

Brief Assessment

Spatial-Frequency Relationships[69] focuses on deepfake detection in facial videos using frequency-domain analysis and attention mechanisms for classification tasks, not resolution-adaptive manipulation localization frameworks for arbitrary images and videos.

3. An Experimental Network Analysis-based Approach for Detection of Jamming Attacks in Wireless Sensor Networks

URL: View paper

Brief Assessment

Jamming Attack Detection[66] addresses jamming attack detection in wireless sensor networks using network metrics (PDR, SNR, PER), which is an entirely different domain from image/video manipulation localization. No technical overlap exists.

4. The detection optimization of low-quality fake face images: feature enhancement and noise suppression strategies

URL: View paper

Brief Assessment

Low-Quality Feature Enhancement[74] focuses on deepfake detection in low-quality facial images using YOLOv9-ARC with adaptive kernel convolution, not on resolution-adaptive manipulation localization frameworks for general images and videos.

5. Deepfake Detection via Spatial-Frequency Attention Network

URL: View paper

Brief Assessment

Spatial-Frequency Attention[72] focuses on deepfake detection using frequency-domain features for face forgery detection, not resolution-adaptive manipulation localization frameworks for general images and videos.

6. TinyDF: Tiny and Effective Model for Deepfake Detection

URL: View paper

Brief Assessment

TinyDF[70] focuses on efficient deepfake detection through lightweight architectures and fusion modules, not on resolution-adaptive frameworks for manipulation localization across arbitrary resolutions without interpolation.

7. Spatial and frequency feature fusion using multi-scale cross attention for enhancing deepfake face detection: M. Uddin et al.

URL: View paper

Brief Assessment

Multi-scale Cross Attention[67] focuses on deepfake face detection using frequency-domain features and multi-scale cross-attention for synthetic face classification, not resolution-adaptive manipulation localization frameworks for arbitrary-resolution images and videos.

8. High-resolution network-based multi-feature fusion for generalized forgery detection

URL: View paper

Brief Assessment

High-resolution Multi-Feature Fusion[75] focuses on multi-resolution feature fusion for forgery detection, not on resolution-adaptive frameworks that process arbitrary resolutions without interpolation or unified image-video modeling architectures.

9. CCM-Net: image splicing localization network based on context-aware and cross-domain multi-scale fusion

URL: View paper

Brief Assessment

CCM-Net[73] focuses on image splicing localization using cross-domain multi-scale fusion and context-aware mechanisms. The candidate does not address resolution-adaptive processing without interpolation or unified image-video manipulation localization, which are the core novelties of RelayFormer.

10. Bznet: Unsupervised multi-scale branch zooming network for detecting low-quality deepfake videos

URL: View paper

Brief Assessment

Branch Zooming Network[68] focuses on deepfake video detection using super-resolution techniques for low-quality compressed videos, not on resolution-adaptive manipulation localization frameworks that handle arbitrary resolutions without interpolation for both images and videos.

Contribution 2: Global Local Relay (GLR) tokens with relay-based attention mechanism

Description: The authors propose GLR tokens that act as information bottlenecks to efficiently exchange global scene-level cues (such as semantic or temporal consistency) across sub-images while preserving local manipulation artifacts, avoiding the computational cost of dense full-resolution attention.

This contribution was assessed against **5 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Inceptive Visual Representation Learning With Diverse Multi-Head Sparse Attention

URL: View paper

Brief Assessment

Sparse Attention Representation[63] focuses on sparse attention mechanisms for visual representation learning, not on relay-based tokens for propagating global context across image partitions in manipulation localization tasks.

2. Multi-Relation Attention Network for Image Patch Matching

URL: View paper

Brief Assessment

Multi-Relation Patch Matching[65] focuses on learning multiple feature relations for patch matching tasks, not on relay-based attention mechanisms for propagating global context across image partitions in manipulation localization.

3. Multi-Disease Detection in Retinal Imaging Using Patch-Based Attention Mechanism

URL: View paper

Brief Assessment

Patch-Based Retinal Detection[64] focuses on medical image classification using patch-based attention for retinal disease detection, not manipulation localization. The attention mechanism operates on ordered image patches for disease classification rather than propagating global context across sub-images for forensic analysis.

4. Conditional diffusion to enhance performance of object detection in unbalanced data engineering drawings

URL: View paper

Brief Assessment

Conditional Diffusion Detection[61] focuses on conditional diffusion models for object detection in unbalanced engineering drawings. The retrieved context mentions 'relay-based railway interlocking systems' and 'image generator', which are unrelated to the relay-based attention mechanism for visual manipulation localization proposed in the original paper.

5. End-to-end object detection with neural networks

URL: View paper

Brief Assessment

End-to-end Neural Detection[62] focuses on object detection using transformers with learnable object queries, not on relay-based attention mechanisms for propagating global context across image partitions in manipulation localization tasks.

Contribution 3: Query-based mask decoder for efficient localization

Description: The authors design a lightweight query-based Transformer decoder that avoids computational bottlenecks by using learnable queries to interact with projected feature maps, enabling efficient mask prediction without excessive overhead.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Mask DINO: Towards A Unified Transformer-based Framework for Object Detection and Segmentation

URL: View paper

Prior Art Analysis

Mask DINO[52] demonstrates prior work on query-based transformer decoders for efficient mask prediction. The candidate paper explicitly describes using learnable query embeddings that interact with projected feature maps to predict masks, which is the same core mechanism claimed in the original contribution. Both papers use content query embeddings to dot-product with pixel embedding maps for mask prediction, and both emphasize efficiency through this query-based approach. The candidate's design predates the original submission and shows that this query-based mask decoder architecture was already established in the literature.

Evidence

Evidence 1 - **Rationale**: Both papers emphasize the lightweight and efficient nature of their query-based mask decoder design, indicating similar design goals and implementation approaches. - **Original**: we design a carefully designed lightweight mask decoder efficiently produces the prediction masks - **Candidate**: this segmentation branch is conceptually simple and easy to implement in the dino framework, as shown in fig. 1.

Evidence 2 - **Rationale**: Both papers describe using learnable queries that interact with feature maps through attention mechanisms for mask prediction, demonstrating the same architectural pattern. - **Original**: a small set of learnable queriesq∈r mf xd then interacts with the projected feature map. the decoder is composed ofkstacked layers. at thek-th layer (k= 1, . . . , k), query features are updated via a cross-attention followed by a self-attention operation - **Candidate**: as dino is not designed for pixel-level alignment as its positional queries are formulated as anchor boxes and its content queries are used to predict box offset and class membership. to perform mask classification, we adopt a key idea from mask2former [3] to construct a pixel embedding map

2. FastInst: A Simple Query-Based Model for Real-Time Instance Segmentation

URL: View paper

Brief Assessment

FastInst[58] focuses on instance segmentation with a query-based decoder for mask prediction, while the original paper addresses visual manipulation localization with a different architectural goal and task domain.

3. An Efficient and Effective Transformer Decoder-Based Framework for Multi-Task Visual Grounding

URL: View paper

Brief Assessment

Transformer Decoder Grounding[59] uses transformer decoder for visual-linguistic fusion with visual features as queries, while the original paper's query-based mask decoder uses learnable queries to interact with projected feature maps for mask prediction—these serve different purposes in distinct architectural contexts.

4. Mask transfiner for high-quality instance segmentation

URL: View paper

Brief Assessment

Mask Transfiner[55] uses learnable queries in a transformer decoder for mask refinement, but operates on sparse incoherent regions detected via a quadtree structure rather than general mask prediction. The original paper's decoder interacts with projected feature maps for manipulation localization, while Mask Transfiner[55] focuses on refining instance segmentation masks in error-prone regions.

5. Query refinement transformer for 3d instance segmentation

URL: View paper

Brief Assessment

Query Refinement Transformer[53] focuses on 3D instance segmentation with learnable queries for mask prediction, while the original paper addresses 2D visual manipulation localization with a different architectural design and task objective.

6. Rethinking Query-Based Transformer for Continual Image Segmentation

URL: View paper

Brief Assessment

Query-Based Continual Segmentation[51] focuses on continual learning for image segmentation with query-based transformers to mitigate catastrophic forgetting, not on designing lightweight decoders for efficient mask prediction in manipulation localization tasks.

7. MGQFormer: Mask-Guided Query-Based Transformer for Image Manipulation Localization

URL: View paper

Brief Assessment

MGQFormer[54] focuses on image manipulation localization using learnable query tokens for mask prediction, while the original paper addresses visual manipulation localization across arbitrary resolutions and video sequences with a relay-based attention mechanism. The technical approaches and problem scopes differ substantially.

8. Mp-former: Mask-piloted transformer for image segmentation

URL: View paper

Brief Assessment

Mp-former[60] focuses on mask-piloted training for image segmentation using ground-truth masks in masked-attention, while the original paper addresses visual manipulation localization with learnable queries for global-local reasoning. The technical approaches and application domains differ substantially.

9. Multi-scale query-based transformer for image forgery localization

URL: View paper

Brief Assessment

Multi-scale Query Transformer[57] focuses on image forgery localization using query-based decoders, while the original paper addresses visual manipulation localization across both images and videos with resolution adaptivity. The technical contexts differ substantially.

10. An Effective Masked Transformer Model for Automatic Modulation Recognition

URL: View paper

Brief Assessment

Masked Transformer Modulation[56] applies masked multi-headed attention in a transformer decoder for modulation recognition tasks, not for visual manipulation localization or mask prediction as in the original paper.

Appendix: Text Similarity Detection

No high-similarity text segments were detected across any compared papers.

References

- [0] RelayFormer: A Unified Local-Global Attention Framework for Scalable Image and Video Manipulation Localization View paper
- [1] A study on content tampering in multimedia watermarking View paper
- [2] Detection of image tampering using multiscale fusion and anomalousness assessment View paper
- [3] Face Forgery Detection via Multi-Feature Fusion and Local Enhancement View paper
- [4] Deepfake Detection Using Convolutional Neural Networks and LSTM Modelling View paper
- [5] Image tamper detection and self-recovery using multiple median watermarking View paper
- [6] Chaotic watermarking for tamper detection: Enhancing robustness and security in digital multimedia View paper
- [7] Towards Analysis Detection of Deepfake Video via Deep Learning Models: A Review View paper
- [8] Fighting fake visual media: a study of current and emerging methods for detecting image and video tampering View paper
- [9] Unravelling digital forgeries: A systematic survey on image manipulation detection and localization View paper
- [10] A novel blind tamper detection and localization scheme for multiple faces in digital images View paper
- [11] A detailed analysis of image and video forgery detection techniques View paper
- [12] An efficient deepfake video detection using robust deep learning View paper
- [13] Visualizing the truth: a survey of multimedia forensic analysis View paper
- [14] An exhaustive review of authentication, tamper detection with localization and recovery techniques for medical images View paper
- [15] An overview of video tampering detection techniques: state-of-the-art and future directions View paper
- [16] Improving Generalization in Facial Manipulation Detection Using Image Noise Residuals and Temporal Features View paper
- [17] Forgery Detection Scheme of Deep Video Frame-rate Up-conversion Based on Dual-stream Multi-scale Spatial-temporal Representation View paper
- [18] An interpretable image tampering detection approach based on cooperative game View paper
- [19] A Comprehensive Survey on Detection of Fake Multimedia Content View paper
- [20] Video tamper detection based on multi-scale mutual information View paper
- [21] Blind Semi-fragile Hybrid Domain-Based Dual Watermarking System for Video Authentication and Tampering Localization View paper
- [22] SRTNet: a spatial and residual based two-stream neural network for deepfakes detection View paper
- [23] VizDefender: Unmasking Visualization Tampering through Proactive Localization and Intent Inference View paper
- [24] A deep neural network with hybrid spotted hyena optimizer and grasshopper optimization algorithm for copy move forgery detection View paper
- [25] An efficient deepfake detection using robust deep learning approch View paper
- [26] Frequency-spatial feature integration with boundary-aware learning for image tampering detection View paper
- [27] Image Tampering Detection View paper
- [28] A novel chaos-based fragile watermarking for image tampering detection and self-recovery View paper
- [29] Image Manipulation Localization Using Attentional Cross-Domain CNN Features View paper
- [30] Tampering Detection and Segmentation Model for Multimedia Forensic View paper
- [31] Image watermarking for tamper detection View paper
- [32] Detecting DeepFake, FaceSwap and Face2Face facial forgeries using frequency CNN View paper
- [33] Multimedia tamper detection and localization using digital watermarking View paper
- [34] Robust image alignment for tampering detection View paper
- [35] Local tampering detection in video sequences View paper
- [36] A hybrid-Sudoku based fragile watermarking scheme for image tampering detection View paper
- [37] A dual-embedded tamper detection framework based on block truncation coding for intelligent multimedia systems View paper
- [38] Tampering detection in compressed digital video using watermarking View paper
- [39] Tampering Detection of Audio-Visual Content Using Encrypted Watermarks View paper
- [40] Large scale image tamper detection and restoration View paper
- [41] News authentication and tampered images: evaluating the photo-truth impact through image verification algorithms View paper
- [42] Malicious inter-frame video tampering detection in MPEG videos using time and spatial domain analysis of quantization effects View paper
- [43] Digital video tampering detection: An overview of passive techniques View paper
- [44] VisGuard: Securing Visualization Dissemination through Tamper-Resistant Data Retrieval View paper
- [45] TAMPAR: Visual Tampering Detection for Parcel Logistics in Postal Supply Chains View paper
- [46] Securing Digital Media Integrity: A Survey of Watermarking and Manipulation Detection for Image Authentication View paper
- [47] Image integrity and tampering detection: A hybrid approach to copy-paste forgery detection using ORB-SSD and CNN View paper
- [48] Deep-Fake Visual Detection using AI View paper
- [49] Video-tampering detection and content reconstruction via self-embedding View paper
- [50] Camera Tampering Detection using Generative Reference Model and Deep Learned Features. View paper
- [51] Rethinking Query-Based Transformer for Continual Image Segmentation View paper
- [52] Mask DINO: Towards A Unified Transformer-based Framework for Object Detection and Segmentation View paper
- [53] Query refinement transformer for 3d instance segmentation View paper
- [54] MGQFormer: Mask-Guided Query-Based Transformer for Image Manipulation Localization View paper
- [55] Mask transfiner for high-quality instance segmentation View paper
- [56] An Effective Masked Transformer Model for Automatic Modulation Recognition View paper
- [57] Multi-scale query-based transformer for image forgery localization View paper
- [58] FastInst: A Simple Query-Based Model for Real-Time Instance Segmentation View paper
- [59] An Efficient and Effective Transformer Decoder-Based Framework for Multi-Task Visual Grounding View paper
- [60] Mp-former: Mask-piloted transformer for image segmentation View paper
- [61] Conditional diffusion to enhance performance of object detection in unbalanced data engineering drawings View paper
- [62] End-to-end object detection with neural networks View paper

- [63] Inceptive Visual Representation Learning With Diverse Multi-Head Sparse Attention View paper
- [64] Multi-Disease Detection in Retinal Imaging Using Patch-Based Attention Mechanism View paper
- [65] Multi-Relation Attention Network for Image Patch Matching View paper
- [66] An Experimental Network Analysis-based Approach for Detection of Jamming Attacks in Wireless Sensor Networks View paper
- [67] Spatial and frequency feature fusion using multi-scale cross attention for enhancing deepfake face detection: M. Uddin et al. View paper
- [68] Bznet: Unsupervised multi-scale branch zooming network for detecting low-quality deepfake videos View paper
- [69] Refining localized attention features with multi-scale relationships for enhanced deepfake detection in spatial-frequency domain View paper
- [70] TinyDF: Tiny and Effective Model for Deepfake Detection View paper
- [71] MSER-Net: Multi-stage edge refinement network for deepfake detection View paper
- [72] Deepfake Detection via Spatial-Frequency Attention Network View paper
- [73] CCM-Net: image splicing localization network based on context-aware and cross-domain multi-scale fusion View paper
- [74] The detection optimization of low-quality fake face images: feature enhancement and noise suppression strategies View paper
- [75] High-resolution network-based multi-feature fusion for generalized forgery detection View paper