Bảng các thuật ngữ:

|  |  |  |  |
| --- | --- | --- | --- |
| **STT** | **Thuật ngữ** | **Thuật ngữ Tiếng Việt** | **Mô tả** |
| 1 | Two-Stream Framework | Khung hai luồng | Khung hai luồng, một kiến trúc sử dụng hai luồng xử lý song song để phân tích dữ liệu. |
| 2 | Spatiotemporal U-Net (STU-Net) | Mạng U-Net không gian- thời gian | Mạng U-Net không gian-thời gian, dùng để dự đoán khung hình tương lai từ một đoạn video. |
| 3 | Context Recovery Stream |  | Luồng phục hồi ngữ cảnh, xử lý dựa trên phục hồi khung hình đầu vào và tái tạo thông tin để dự đoán khung hình tương lai. |
| 4 | Maximum Local Error (MLE) |  | Lỗi cục bộ lớn nhất, cơ chế tập trung vào sự khác biệt trong khu vực bất thường để cải thiện độ chính xác của điểm số bất thường. |
| 5 | Improved Learnable Locality-Sensitive Hashing (iL2SH) |  | Phương pháp băm nhạy cảm cục bộ có thể học được cải tiến, dùng để lưu trữ và truy xuất kiến thức về trạng thái bình thường. |
| 6 | Knowledge Retrieval Stream |  | Luồng truy xuất tri thức, dựa trên việc lấy lại thông tin quan trọng từ dữ liệu để đánh giá tính bất thường. |
| 7 | Knowledge Base |  | Cơ sở tri thức, lưu trữ thông tin về tính bình thường được trích xuất từ dữ liệu huấn luyện. |
| 8 | Anomaly Score |  | Điểm số đánh giá mức độ bất thường tại một thời điểm nhất định. |
| 9 | Encoder |  | Bộ mã hóa, trích xuất các đặc trưng ngữ nghĩa cấp cao từ đầu vào. |
| 10 | Decoder |  | Bộ giải mã, khôi phục bản đồ đặc trưng bằng cách tăng dần độ phân giải không gian. |
| 11 | Shortcut |  | Kết nối giữa các cấp của bộ mã hóa và bộ giải mã nhằm tránh mất mát gradient. |
| 12 | Temporal Squeezing Layer (TSL) |  | Lớp nén thời gian, kết hợp các đặc trưng từ các bước thời gian khác nhau và nén chiều thời gian thành 1. |
| 13 | I3D Network |  | Mạng tích chập 3D được áp dụng trong bộ mã hóa. |
| 14 | ResNet-50 |  | Một mô hình mạng sâu làm xương sống cho bộ mã hóa, huấn luyện trước trên ImageNet. |
| 15 | Residual Connection |  | Kết nối dư, phương pháp trong mạng sâu để tránh mất mát thông tin khi đi qua các lớp mạng. |
| 16 | Batch Normalization (BN) |  | Chuẩn hóa batch, giúp tối ưu hóa quá trình huấn luyện. |
| 17 | ReLU và Leaky ReLU |  | Các hàm kích hoạt, trong đó Leaky ReLU giúp tránh tình trạng chết gradient. |
| 18 | Transposed Convolution |  | Tích chập chuyển vị, mở rộng độ phân giải của bản đồ đặc trưng. |
| 19 | Mean Square Error (MSE) và L1 Loss |  | Các hàm lỗi phổ biến trong huấn luyện mạng. |
| 20 | Frame-Level Error (FLE) |  | Lỗi ở mức khung hình, dùng làm thước đo phát hiện bất thường. |
| 21 | Optical Flow |  | Phương pháp bổ sung thông tin chuyển động giữa các khung hình. |
| 22 | Data Augmentation |  | Tăng cường dữ liệu, cải thiện bộ dữ liệu bằng cách chỉnh sửa video ban đầu. |
| 23 | Hash Function |  | Hàm băm, ánh xạ các đặc trưng vào không gian mã hóa. |
| 24 | Bucket |  | Phân vùng trong bảng băm. |
| 25 | Binary Key |  | Khóa nhị phân, được sử dụng trong truy xuất dữ liệu. |
| 26 | L2 Distance |  | Khoảng cách L2, đo lường sự khác biệt giữa hai vector. |
| 27 | Anomaly Detection (VAD) |  | Phát hiện bất thường trong video, dựa trên phân tích dữ liệu không gian-thời gian. |
| 28 | AUC (Area Under Curve) |  | Diện tích dưới đường cong, thước đo hiệu suất của mô hình. |
| 29 | ROC (Receiver Operating Characteristic) |  | Đường cong ROC, biểu diễn quan hệ giữa tỷ lệ dương tính thật và dương tính giả. |
| 30 | Ablation Study |  | Nghiên cứu cắt giảm để phân tích đóng góp của từng thành phần trong mô hình. |
| 31 | Siamese Network |  | Mạng nơ-ron kiểu Siamese, sử dụng các nhánh song song để so sánh đặc trưng giữa các mẫu. |
| 32 | Memory Modules |  | Các mô-đun bộ nhớ, lưu trữ mẫu hình bình thường trong khôi phục ngữ cảnh. |
| 33 | Temporal Sampling Rate |  | Tốc độ lấy mẫu theo thời gian. |
| 34 | Negative Samples |  | Các mẫu âm, dữ liệu không thuộc về lớp mục tiêu. |
| 35 | Gaussian Filter |  | Bộ lọc Gaussian, làm mịn dữ liệu không gian hoặc thời gian. |
| 36 | Sliding Window |  | Cửa sổ trượt, một phương pháp di chuyển cửa sổ qua dữ liệu để phân tích từng đoạn nhỏ. |
| 37 | Fusion |  | Hợp nhất kết quả từ các dòng xử lý. |
| 38 | Late Fusion |  | Hợp nhất trễ, kết hợp kết quả ở giai đoạn sau của xử lý dữ liệu. |
| 39 | Micro-AUC |  | AUC tính trên toàn bộ dữ liệu chi tiết. |
| 40 | Macro-AUC |  | AUC trung bình trên từng video. |
| 41 | Event Encoder |  | Bộ mã hóa sự kiện, xử lý thông tin sự kiện trong video. |
| 42 | Compact Long Hash Code |  | Mã băm dài gọn, đại diện dữ liệu trong không gian mã hóa. |
| 43 | Cluster Attention Module |  | Mô-đun chú ý cụm, ánh xạ các đặc trưng vào không gian cụm để phân tích. |
| 44 | Exemplar Set |  | Tập hợp mẫu điển hình, lưu trữ các mẫu đại diện của dữ liệu bình thường. |
| 45 | Mutual Difference Loss |  | Hàm mất mát độ khác biệt tương hỗ, tối ưu hóa sự khác biệt giữa các mẫu. |

Phần III

1. **Two-Stream Framework (Khung Hai Luồng)**: Một kiến trúc sử dụng hai luồng xử lý song song.
2. **Spatiotemporal U-Net (STU-Net)**: Mạng U-Net không gian-thời gian, được dùng để dự đoán khung hình tương lai.
3. **Context Recovery Stream (Luồng Phục Hồi Ngữ Cảnh)**: Luồng xử lý dựa trên phục hồi khung hình đầu vào.
4. **Maximum Local Error (MLE)**: Cơ chế lỗi cục bộ lớn nhất, được sử dụng để đánh giá mức độ bất thường.
5. **Improved Learnable Locality-Sensitive Hashing (iL2SH)**: Băm nhạy cảm cục bộ có thể học được cải tiến, dùng cho truy xuất tri thức.
6. **Knowledge Retrieval Stream (Luồng Truy Xuất Tri Thức)**: Luồng xử lý dựa trên việc truy xuất thông tin về tính bình thường từ cơ sở tri thức.
7. **Knowledge Base (Cơ Sở Tri Thức)**: Lưu trữ thông tin về tính bình thường được trích xuất từ dữ liệu huấn luyện.
8. **Anomaly Score (Điểm Bất Thường)**: Điểm số đánh giá mức độ bất thường tại một thời điểm τ.
9. **Spatiotemporal U-Net (STU-Net)**: Mạng U-Net không gian-thời gian, dùng để dự đoán khung hình tương lai từ một đoạn video.
10. **Context Recovery Stream (Luồng Phục Hồi Ngữ Cảnh)**: Một luồng xử lý trong khung hai luồng, dùng để phục hồi ngữ cảnh của video.
11. **Maximum Local Error (MLE)**: Lỗi cục bộ lớn nhất, dùng để đánh giá mức độ bất thường.
12. **Encoder (Bộ Mã Hóa)**: Phần mạng trích xuất các đặc trưng ngữ nghĩa cấp cao từ đầu vào.
13. **Decoder (Bộ Giải Mã)**: Phần mạng khôi phục bản đồ đặc trưng bằng cách tăng dần độ phân giải không gian.
14. **Shortcut**: Kết nối giữa các cấp của bộ mã hóa và bộ giải mã nhằm tránh mất mát gradient.
15. **Temporal Squeezing Layer (TSL)**: Lớp nén thời gian, kết hợp các đặc trưng từ các bước thời gian khác nhau và nén chiều thời gian thành 1.
16. **I3D Network**: Một mạng tích chập 3D được áp dụng trong bộ mã hóa.
17. **ResNet-50**: Một mô hình mạng sâu làm xương sống cho bộ mã hóa.
18. **Residual Connection (Kết Nối Dư)**: Một phương pháp trong mạng sâu để tránh mất mát thông tin khi đi qua các lớp mạng.
19. **Batch Normalization (BN)**: Chuẩn hóa batch, giúp tối ưu hóa quá trình huấn luyện.
20. **ReLU** và **Leaky ReLU**: Các hàm kích hoạt, trong đó Leaky ReLU giúp tránh tình trạng chết gradient.
21. **Transposed Convolution**: Tích chập chuyển vị, mở rộng độ phân giải của bản đồ đặc trưng.
22. **Temporal Squeezing Layer (TSL)**: Lớp nén thời gian, nén chiều thời gian của bản đồ đặc trưng.
23. **Mean Square Error (MSE)** và **L1 Loss**: Các hàm lỗi phổ biến trong huấn luyện mạng.
24. **Frame-Level Error (FLE)**: Lỗi ở mức khung hình, dùng làm thước đo phát hiện bất thường.
25. **Optical Flow**: Phương pháp bổ sung thông tin chuyển động giữa các khung hình.
26. **Maximum Local Error (MLE)**: Phương pháp tập trung vào lỗi lớn nhất trong vùng bất thường.
27. **Max Pooling**: Một kỹ thuật tính toán giá trị lớn nhất trong cửa sổ trượt.
28. **Data Augmentation**: Tăng cường dữ liệu, phương pháp cải thiện bộ dữ liệu bằng cách chỉnh sửa video ban đầu.
29. **Knowledge Retrieval Stream**: Dòng truy xuất tri thức.
30. **iL2SH (improved learnable locality-sensitive hashing)**: Phương pháp băm nhạy cảm cục bộ có thể học được cải tiến.
31. **Hash function**: Hàm băm.
32. **Bucket**: Phân vùng/bucket trong bảng băm.
33. **Event encoder**: Bộ mã hóa sự kiện.
34. **Hash encoder**: Bộ mã hóa băm.
35. **Compact long hash code**: Mã băm dài gọn.
36. **Cosine distance**: Khoảng cách cosine.
37. **Mutual difference loss**: Mất mát độ khác biệt tương hỗ.
38. **Siamese network**: Mạng Siamese.
39. **Snippet**: Đoạn video.
40. **Layer normalization**: Chuẩn hóa lớp.
41. **Sigmoid activation**: Hàm kích hoạt sigmoid.
42. **Knowledge base**: Cơ sở tri thức.
43. **Hash table**: Bảng băm.
44. **Bucket**: Phân vùng trong bảng băm.
45. **Key-value pair**: Cặp khóa-giá trị.
46. **Binary key**: Khóa nhị phân.
47. **Mean vector**: Vector trung bình.
48. **L2 distance**: Khoảng cách L2.
49. **Anomaly score**: Điểm bất thường.
50. **Pmax**: Điểm bất thường cao nhất.
51. **Fusion**: Hợp nhất kết quả từ các dòng.
52. **Late fusion**: Hợp nhất trễ (kết hợp kết quả ở giai đoạn sau).
53. **Context recovery stream**: Dòng phục hồi ngữ cảnh.
54. **Knowledge retrieval stream**: Dòng truy xuất tri thức.
55. **VAD (Video Anomaly Detection)**: Phát hiện bất thường trong video.
56. **Two-Stream Framework**: Mô hình hai luồng.
57. **AUC (Area Under Curve)**: Diện tích dưới đường cong.
58. **ROC (Receiver Operating Characteristic)**: Đường cong ROC.
59. **Micro-AUC**: AUC tính trên toàn bộ dữ liệu.
60. **Macro-AUC**: AUC trung bình trên từng video.
61. **STU-Net**: Mạng học sâu sử dụng để khôi phục ngữ cảnh.
62. **Sliding Window**: Cửa sổ trượt.
63. **Gaussian Filter**: Bộ lọc Gaussian.
64. **Hash Encoder**: Bộ mã hóa băm.
65. **I3D (Inflated 3D Network)**: Mạng học sâu tiền huấn luyện trên dữ liệu động.
66. **Ablation Study**: Nghiên cứu cắt giảm (để phân tích đóng góp của từng thành phần trong mô hình).
67. **Two-stream Framework**: Khung làm việc hai luồng (phân chia thành hai hướng xử lý dữ liệu chính).
68. **Context Recovery Stream**: Luồng khôi phục ngữ cảnh (tập trung tái tạo ngữ cảnh).
69. **STU-Net**: Một biến thể của U-Net có thể tận dụng chuyển động của đầu vào.
70. **ResNet-50**: Mạng học sâu dùng làm bộ mã hóa (Encoder), huấn luyện trước trên ImageNet.
71. **Knowledge Retrieval Stream**: Luồng truy xuất kiến thức (tập trung vào việc lấy lại thông tin quan trọng từ dữ liệu).
72. **iL2SH**: Một mô hình hash (được sử dụng trong luồng truy xuất kiến thức).
73. **w/ neg.**: Được huấn luyện với các cặp âm (negative pairs).
74. **w/ Lm**: Được huấn luyện với hàm mất mát chênh lệch tương hỗ (mutual difference loss).
75. **MLE (Motion Level Encoder)**: Bộ mã hóa mức độ chuyển động.
76. **Micro-AUC**: Diện tích dưới đường cong ROC, tính trên toàn bộ khung hình của tập dữ liệu.
77. **Macro-AUC**: Trung bình của AUC trên các video.
78. **LLSH**: Locality-Sensitive Hashing, một phương pháp xử lý dữ liệu khác được so sánh.
79. **CR (Context Recovery):** Phục hồi ngữ cảnh.
80. **KR (Knowledge Retrieval):** Truy xuất tri thức.
81. **AUC (Area Under Curve):** Diện tích dưới đường cong, chỉ số đánh giá hiệu năng.
82. **Fusion Study:** Nghiên cứu hợp nhất.
83. **STU-Net:** Mạng nơ-ron dùng để phục hồi ngữ cảnh.
84. **iL2SH:** Mô hình truy xuất tri thức.
85. **Micro-AUC:** Chỉ số AUC ở mức vi mô.
86. **Temporal Sampling Rate:** Tốc độ lấy mẫu theo thời gian.
87. **Memory-Augmented Models:** Mô hình có bộ nhớ tăng cường.
88. **Maximum Local Error (MLE)**: Lỗi cục bộ tối đa, một phương pháp phát hiện bất thường bằng cách sử dụng cửa sổ trượt.
89. **Frame-Level Error (FLE)**: Lỗi ở mức khung hình.
90. **Score Gaps**: Khoảng cách điểm, thể hiện sự khác biệt giữa điểm trung bình của khung hình bất thường và bình thường.
91. **AUC (Area Under Curve)**: Diện tích dưới đường cong, thước đo hiệu suất của mô hình.
92. **Fusion of Two Streams**: Sự hợp nhất của hai luồng, kết hợp đầu ra của hai phương pháp để cải thiện hiệu suất phát hiện bất thường.
93. **micro-AUC**: Diện tích dưới đường cong vi mô, một thước đo hiệu suất mô hình trên cấp độ chi tiết.
94. **macro-AUC**: Diện tích dưới đường cong vĩ mô, một thước đo hiệu suất mô hình trên cấp độ tổng quan.
95. **LLSH**: Phương pháp phát hiện bất thường sử dụng hashing.
96. **object detection (w/ obj.)**: Phát hiện đối tượng, một kỹ thuật sử dụng để nhận dạng và định vị các đối tượng trong video.
97. **pose estimation**: Đánh giá tư thế, một kỹ thuật để xác định vị trí và hướng của cơ thể hoặc các bộ phận của cơ thể trong video.
98. **optical flow**: Dòng quang học, một phương pháp phân tích chuyển động giữa các khung hình liên tiếp trong video.
99. **two-stream framework**: Khung hai luồng, một kiến trúc mạng sử dụng hai luồng xử lý độc lập để giải quyết vấn đề.
100. **context recovery stream**: Luồng phục hồi ngữ cảnh, luồng xử lý để khôi phục thông tin từ video để dự đoán các khung hình tương lai.
101. **knowledge retrieval stream**: Luồng truy xuất kiến thức, luồng xử lý để lưu trữ và truy xuất kiến thức về trạng thái bình thường nhằm phát hiện bất thường.
102. **spatiotemporal U-Net (STU-Net)**: Mạng U-Net không gian-thời gian, một phiên bản mở rộng của U-Net để xử lý dữ liệu không gian-thời gian.
103. **maximum local error (MLE)**: Lỗi cục bộ tối đa, một cơ chế tập trung vào sự khác biệt trong khu vực bất thường để cải thiện độ chính xác của điểm số bất thường.
104. **improved learnable locality-sensitive hashing (iL2SH)**: Phương pháp băm nhạy cảm theo địa phương có thể học được cải tiến, được sử dụng để lưu trữ và truy xuất kiến thức về trạng thái bình thường.
105. **OC-SVM (One-Class Support Vector Machine):** Máy vector hỗ trợ cho bài toán phân lớp một lớp.
106. **Nearest Neighbor Samples:** Các mẫu gần nhất.
107. **Cluster Centers:** Tâm cụm, đại diện cho trung tâm của các cụm dữ liệu.
108. **Probability Distributions:** Phân phối xác suất.
109. **K-means:** Thuật toán phân cụm dữ liệu thành **K cụm**.
110. **One-Versus-Rest SVMs:** Phương pháp phân lớp sử dụng nhiều bộ phân loại SVM, mỗi bộ phân loại đối lập một lớp với tất cả các lớp khác.
111. **Cluster Attention Module:** Mô-đun chú ý cụm, ánh xạ các đặc trưng vào không gian cụm.
112. **Exemplar Set:** Tập hợp mẫu điển hình, lưu trữ các mẫu đại diện của dữ liệu bình thường.
113. **MoCo (Momentum Contrast):** Kỹ thuật học sâu dùng để huấn luyện các hàm băm dựa trên đối lập động lượng.
114. **Negative Samples:** Các mẫu âm, dữ liệu không thuộc về lớp mục tiêu.
115. **Siamese Network:** Mạng nơ-ron kiểu Siamese, sử dụng các nhánh song song để so sánh đặc trưng giữa các mẫu.
116. **Hash Functions:** Hàm băm, ánh xạ các đặc trưng vào không gian mã hóa.
117. **Memory Modules:** Các mô-đun bộ nhớ, lưu trữ mẫu hình bình thường trong khôi phục ngữ cảnh.