A Benchmark on Multimodal Deepfake Continual Detection

A Baselines Desription

We implement and adapt existing continual learning algorithms to MDCD scenarios. The strategies they adopted are illustrated in Table 4, which are primarily divided into four types. Exemplar represents the method that preserves a certain number of exemplars for replay. Distillation involves adding a constraint term to the model update. Prototype denotes the method that retains old knowledge by calculating a class prototype. Analytic indicates that the method leverages analytic learning algorithms to update weights of the classification head. Each baseline is described detailly as follows:

- Finetune: training on each incremental dataset directly. This method easily leads to catastrophic forgetting of historical deepfake knowledge due to domain gaps.
- Replay: randomly preserves some examples from each trained dataset and re-trains with them in the subsequent tasks. It mitigates catastrophic forgetting by preserving previous training samples.
- iCaRL (CVPR 2017): selects the samples closest to the class mean at each phase and replays them in the following.
- LwF (TPAMI 2018): uses only new task data to train the network while preserving the original capabilities by introducing a regularization loss between the current network and the old network.
- PASS (CVPR 2021): augments the embedding of the prototype of the old class in the feature space to retain the old decision boundary, and reduce task-level overfitting through self-supervised learning. It is an exemplar-free method without extra storage cost.
- ACIL (NeurIPS 2022): trains the feature extractor in the base training phase and updates the analytic classifier with the frozen feature extractor in incremental learning phases. ACIL is an AL-based paradigm using least squares to calculate the closed-form solution of the classification weight. It achieves a comparable performance to joint learning on all seen datasets.
- Fetril (WACV 2023): generates pseudo-features of old classes by simply translating the features of new classes with the stored old class mean. It only updates the classifier rather than the whole network, gaining a lower time cost.
- DFIL (MM 2023): selects central and hard representative samples for replay to cope with different deepfake data distributions. It regularizes the model in both feature-level and label-level to mitigate catastrophic forgetting. DFIL is the first domain-incremental learning framework for deepfake detection, and it is a model-agnostic method that easily adapts to MDCD.
- DS-AL (AAAI 2024): introduces a compensation stream improving the inherent under-fitting limitation due to adopting linear mapping based on ACIL.
- MMAL (MM 2024): is a multimodal analytic learning framework introducing a modality fusion module that harnesses individual knowledge from audio and visual modality in

Table 4: The strategies that baselines used.

Method	Exemplar	Distillation	Prototype	Analytic
Finetune				
Replay	√			
iCaRL	√			
LwF		√		
PASS		√	√	
ACIL				√
Fetril			√	
DFIL	√	√		
DS-AL				√
MMAL				√
TagFex	√	√		

tandem. MMAL leverages fine-grained modality knowledge to alleviate the under-fitting limitations.
- TagFex (CVPR 2025): provides additional task-agnostic features for subsequent tasks and integrates the task-agnostic features with the task-specific features. It helps capture diverse features to combat the feature collision.

Note that the memory buffer size for replay-based continual baselines varies from different incremental settings. We set the number of exemplars per class to 500, 100, and 1000 for the sequence incremental setting, cross-technology incremental setting, and cross-modality setting, respectively. The critical hyperparameters of each baseline are referenced from the original algorithm, unless specifically stated.

B Additional Experiments

In order to promote further research in the MDCD scenario, we provide detailed experimental details for thorough evaluation. Tables 5 and 6 respectively detail the performance of each stage in various tasks under the sequential incremental scenario in both forward and reverse orders. We find an unavoidable forgetting for every baseline. Even the most effective methods can preserve forged knowledge well in some tasks, but they still struggle to address the feature differences between multimodal deepfake data, especially for Task 2 in the forward sequence. Multimodal data gaps are challenges in MDCD that need further efforts.

Table 7 illustrates the detailed performance of each phase in various tasks in the modality incremental scenario. Finetune and PASS suffer 33.63% and 34.82% accuracy drops on video-only deepfake videos after training on audio-only deepfake data, while others retain video deepfake knowledge. It demonstrates that the deepfake artifact features between the visual and audio unimodal are relatively simple and direct transformations. Nevertheless, most baselines encounter catastrophic forgetting after training on video-audio deepfakes. Especially, AL-based methods still drop over 16% accuracy, though freezing the feature extractor. thereby verifying a more complex feature drift from unimodal deepfakes to multimodal deepfakes.

Table 5: Detailed performance at each phase in the forward sequence incremental setting.

Method	Datasets	Task1 ACC	Task1 AUC	Task2 ACC	Task2 AUC	Task3 ACC	Task3 AUC	Task4 ACC	Task4 AUC	Task5 ACC	Task5 AUC	Task6 ACC	Task6 AUC	AAUC ↑	Fk ↓
Finetune	Task1	89.94	78.46	-	-	-	-	-	-	-	-	-	-	89.94	-
	Task2	72.85	74.18	83.94	84.83	-	-	-	-	-	-	-	-	78.40	17.09
	Task3	49.22	59.19	54.21	58.13	99.81	99.69	-	-	-	-	-	-	67.75	35.20
	Task4	78.91	52.59	55.49	48.97	69.53	49.90	98.53	86.17	-	-	-	-	59.41	23.25
	Task5	25.00	36.08	44.97	50.81	71.24	52.57	33.41	45.57	98.33	98.75	-	-	56.76	49.40
	Task6	19.73	46.54	42.40	49.85	34.85	41.29	3.37	50.09	28.49	50.00	99.76	99.84	56.27	68.34
Replay	Task1	85.06	85.30	-	-	-	-	-	-	-	-	-	-	85.30	-
	Task2	83.69	80.56	82.77	83.58	-	-	-	-	-	-	-	-	82.07	1.37
	Task3	86.04	77.50	63.38	64.55	99.59	99.35	-	-	-	-	-	-	80.47	9.69
	Task4	86.62	74.16	66.84	66.39	88.58	82.68	97.90	72.34	-	-	-	-	73.89	8.98
	Task5	85.45	65.84	60.33	53.72	95.81	94.64	95.62	50.26	99.29	98.90	-	-	72.67	7.42
	Task6	84.96	66.37	56.94	50.44	93.43	90.42	96.70	97.25	71.20	51.07	99.98	99.97	75.92	12.59
iCaRL	Task1	85.06	85.30	-	-	-	-	-	-	-	-	-	-	85.30	-
	Task2	83.11	69.91	82.74	81.07	-	-	-	-	-	-	-	-	75.49	1.95
	Task3	60.94	74.88	52.78	58.77	98.54	97.58	-	-	-	-	-	-	77.08	27.04
	Task4	86.52	84.37	71.46	71.78	97.02	97.46	95.05	95.70	-	-	-	-	87.33	4.27
	Task5	87.50	89.04	62.79	66.31	97.56	97.37	86.78	88.38	98.79	98.96	-	-	88.01	7.30
	Task6	66.21	77.05	52.71	58.28	50.38	64.44	62.07	80.41	44.40	61.10	97.42	98.27	73.26	37.37
LwF	Task1	85.25	85.42	-	-	-	-	-	-	-	-	-	-	85.42	-
	Task2	54.69	70.75	52.71	58.65	-	-	-	-	-	-	-	-	64.70	30.56
	Task3	72.07	68.64	59.66	60.95	90.15	92.87	-	-	-	-	-	-	74.15	6.59
	Task4	78.81	50.00	57.49	50.04	69.78	50.00	97.32	60.29	-	-	-	-	52.58	9.66
	Task5	25.78	36.74	46.55	51.95	70.51	55.47	30.85	46.42	99.31	99.23	-	-	57.96	39.67
	Task6	21.19	46.63	43.31	50.28	31.06	38.60	3.54	50.18	30.09	51.12	99.95	99.95	56.13	60.50
PASS	Task1	90.33	83.93	-	-	-	-	-	-	-	-	-	-	83.93	-
	Task2	79.88	75.95	62.23	63.34	-	-	-	-	-	-	-	-	69.65	10.45
	Task3	71.58	69.34	52.56	54.43	58.05	45.84	-	-	-	-	-	-	56.54	14.21
	Task4	37.01	50.10	44.08	48.85	52.67	49.06	40.86	44.63	-	-	-	-	48.16	25.62
	Task5	74.22	54.67	59.22	53.01	52.19	49.58	84.63	64.18	71.42	50.13	-	-	54.31	6.25
	Task6	79.20	62.55	59.39	54.52	41.41	53.72	83.94	72.54	71.14	50.37	56.52	50.68	57.40	6.32
ACIL	Task1	85.55	76.85	-	-	-	-	-	-	-	-	-	-	85.55	-
	Task2	83.01	72.54	59.87	58.99	-	-	-	-	-	-	-	-	65.77	2.54
	Task3	81.05	68.61	61.53	60.38	80.93	74.51	-	-	-	-	-	-	67.83	2.25
	Task4	84.96	68.22	63.19	61.79	81.14	74.00	95.33	50.11	-	-	-	-	63.53	0.20
	Task5	85.45	69.54	63.68	62.10	80.60	72.90	95.54	50.22	69.44	51.90	-	-	61.33	0.16
	Task6	82.23	61.43	63.55	61.25	78.71	69.06	96.63	50.35	71.26	50.20	75.07	50.10	57.07	1.18
Fetril	Task1	86.04	85.75	-	-	-	-	-	-	-	-	-	-	85.75	-
	Task2	82.42	69.64	64.48	63.85	-	-	-	-	-	-	-	-	66.75	3.62
	Task3	70.90	54.41	53.91	51.52	78.03	66.65	-	-	-	-	-	-	57.53	12.86
	Task4	78.81	50.00	57.47	50.00	69.81	50.04	96.77	49.99	-	-	-	-	50.12	7.49
	Task5	53.12	36.57	52.57	47.12	61.39	44.75	87.64	51.37	71.34	50.20	-	-	46.36	17.66
	Task6	78.03	49.50	57.39	49.94	69.78	50.00	96.77	49.99	71.38	49.95	75.24	49.99	49.90	4.68
DFIL	Task1	84.38	63.81	-	-	-	-	-	-	-	-	-	-	63.81	-
	Task2	84.96	72.77	82.70	80.50	-	-	-	-	-	-	-	-	76.64	0.00
	Task3	85.64	66.97	67.38	64.50	99.62	99.50	-	-	-	-	-	-	76.99	7.66
	Task4	85.55	69.10	65.34	61.01	94.72	91.75	97.75	66.17	-	-	-	-	72.01	7.45
	Task5	85.06	77.55	62.31	63.58	92.42	94.57	91.71	76.55	99.00	99.20	-	-	82.29	8.55
	Task6	84.86	79.45	57.30	58.00	95.18	93.66	95.54	97.69	61.95	47.03	99.81	99.62	79.24	13.98
DS-AL	Task1	85.55	76.85	-	-	-	-	-	-	-	-	-	-	76.85	-
	Task2	81.64	71.00	59.36	58.45	-	-	-	-	-	-	-	-	64.73	3.91
	Task3	80.27	67.61	61.22	60.17	81.63	75.44	-	-	-	-	-	-	67.74	2.64
	Task4	84.18	67.39	62.73	61.30	81.95	74.91	95.19	49.17	-	-	-	-	63.19	0.46
	Task5	83.98	67.43	63.05	61.32	81.25	73.70	94.96	51.23	68.28	52.25	-	-	61.19	0.63
	Task6	82.23	61.10	63.15	60.69	78.98	69.43	96.43	49.81	71.23	50.42	74.88	50.11	56.93	1.26
MMAL	Task1	89.75	84.23	-	-	-	-	-	-	-	-	-	-	84.23	-
	Task2	79.79	68.31	62.27	61.29	-	-	-	-	-	-	-	-	64.80	9.96
	Task3	78.61	65.54	61.62	60.44	80.90	74.44	-	-	-	-	-	-	66.81	5.90
	Task4	82.81	65.34	62.69	60.54	80.49	73.66	95.07	50.41	-	-	-	-	62.49	2.45
	Task5	82.81	66.35	62.42	60.35	80.33	72.74	92.43	52.09	69.47	55.45	-	-	61.40	2.60
	Task6	82.13	61.71	60.96	56.98	77.00	66.01	94.47	50.54	70.82	52.02	74.19	50.63	56.32	2.77
TagFex	Task1	89.45	79.33	-	-	-	-	-	-	-	-	-	-	79.33	-
	Task2	85.06	65.93	76.73	72.99	-	-	-	-	-	-	-	-	69.46	0.00
	Task3	83.30	62.45	74.43	72.42	99.13	98.64	-	-	-	-	-	-	77.84	2.20
	Task4	86.52	69.55	66.48	62.46	98.16	98.22	97.18	61.52	-	-	-	-	72.94	2.82
	Task5	57.62	64.86	62.75	66.49	49.30	63.39	69.84	65.69	83.92	88.74	-	-	69.83	3.54
	Task6	41.80	62.23	49.24	55.62	92.59	94.18	57.60	78.10	46.44	62.28	99.89	99.90	75.39	24.60

Table 6: Performance comparison of different methods on various datasets reverse

Method	Datasets	Task6 ACC	Task6 AUC	Task5 ACC	Task5 AUC	Task4 ACC	Task4 AUC	Task3 ACC	Task3 AUC	Task2 ACC	Task2 AUC	Task1 ACC	Task1 AUC	AAUC ↑	Fk ↓
Finetune	Task6	98.85	98.90	-	-	-	-	-	-	-	-	-	-	98.85	-
	Task5	26.44	50.13	94.54	93.66	-	-	-	-	-	-	-	-	60.40	72.41
	Task4	75.27	50.00	71.51	50.00	96.80	50.00	-	-	-	-	-	-	81.20	23.31
	Task3	22.86	17.97	68.16	49.67	61.20	34.66	97.08	95.17	-	-	-	-	62.33	45.99
	Task2	72.98	50.48	71.67	50.52	91.33	51.96	50.16	54.85	75.46	75.74	-	-	72.32	25.28
	Task1	52.57	50.67	64.59	51.87	65.96	64.56	43.18	52.92	61.13	64.07	91.41	82.59	63.13	35.06
Replay	Task6	98.85	98.90	-	-	-	-	-	-	-	-	-	-	98.85	-
	Task5	89.89	92.58	97.08	96.71	-	-	-	-	-	-	-	-	93.49	8.96
	Task4	87.80	76.69	76.69	61.91	96.89	74.87	-	-	-	-	-	-	87.13	15.72
	Task3	93.70	87.90	82.34	82.21	73.68	75.95	98.51	98.10	-	-	-	-	87.06	14.37
	Task2	81.19	62.68	73.91	55.25	91.36	59.82	80.90	70.25	81.63	80.23	-	-	81.80	15.99
	Task1	56.45	60.96	67.18	64.94	79.55	72.45	79.55	66.25	70.77	69.75	90.14	78.75	73.93	23.89
iCaRL	Task6	98.85	98.90	-	-	-	-	-	-	-	-	-	-	98.85	-
	Task5	26.49	50.21	94.58	93.73	-	-	-	-	-	-	-	-	60.53	72.36
	Task4	76.02	51.52	72.63	52.88	96.86	50.90	-	-	-	-	-	-	81.84	22.39
	Task3	80.98	85.75	84.99	88.75	77.48	86.19	97.00	96.96	-	-	-	-	85.11	15.61
	Task2	47.76	59.35	76.17	77.23	74.11	78.79	72.94	79.32	78.62	77.73	-	-	69.92	29.08
	Task1	31.37	53.80	64.46	74.66	79.06	89.18	91.15	89.78	67.76	70.27	88.48	87.30	70.38	26.42
LwF	Task6	98.85	98.90	-	-	-	-	-	-	-	-	-	-	98.86	-
	Task5	28.98	52.81	45.02	61.54	-	-	-	-	-	-	-	-	37.00	69.87
	Task4	58.68	60.77	73.35	65.58	52.16	65.71	-	-	-	-	-	-	61.40	20.08
	Task3	50.85	36.56	69.11	50.66	75.00	44.84	90.85	84.95	-	-	-	-	71.45	17.41
	Task2	74.92	50.08	71.54	50.05	96.14	49.66	52.54	52.47	69.53	65.39	-	-	72.93	16.04
	Task1	60.34	50.36	69.31	50.07	77.82	61.54	42.78	56.81	57.64	60.89	91.11	82.06	66.50	24.17
PASS	Task6	98.85	98.90	-	-	-	-	-	-	-	-	-	-	98.85	-
	Task5	24.93	50.06	87.41	91.05	-	-	-	-	-	-	-	-	56.17	73.92
	Task4	25.49	50.06	28.48	49.98	5.58	50.91	-	-	-	-	-	-	19.92	66.15
	Task3	67.49	49.26	12.47	15.87	92.67	50.04	73.70	56.48	-	-	-	-	61.59	35.43
	Task2	62.22	48.62	13.22	21.26	86.24	51.51	74.54	58.04	50.06	46.45	-	-	57.26	29.31
	Task1	27.02	49.03	28.22	49.53	15.17	48.78	67.12	60.15	42.69	46.32	34.86	49.58	35.85	44.66
ACIL	Task6	99.57	99.39	-	-	-	-	-	-	-	-	-	-	99.57	-
	Task5	99.29	98.76	83.76	88.15	-	-	-	-	-	-	-	-	91.52	0.28
	Task4	99.06	98.28	83.93	88.02	53.60	71.24	-	-	-	-	-	-	78.86	0.25
	Task3	98.89	97.95	83.93	87.96	54.44	72.11	85.17	76.92	-	-	-	-	80.61	0.23
	Task2	98.76	97.68	83.96	87.80	54.81	71.43	85.36	77.08	50.86	49.78	-	-	74.75	0.20
	Task1	98.71	97.57	83.88	87.68	54.98	71.96	85.36	76.90	51.94	49.55	74.71	52.45	74.93	0.19
Fetril	Task6	98.85	98.90	-	-	-	-	-	-	-	-	-	-	98.85	-
	Task5	92.73	87.62	84.49	83.51	-	-	-	-	-	-	-	-	88.61	6.12
	Task4	75.27	50.00	74.40	49.93	96.80	50.00	-	-	-	-	-	-	81.16	16.83
	Task3	23.64	26.21	67.07	57.62	55.24	45.08	85.82	77.36	-	-	-	-	57.94	44.73
	Task2	23.50	15.70	64.41	45.10	85.02	44.35	73.24	56.03	57.26	50.50	-	-	60.69	29.95
	Task1	19.99	13.36	62.27	43.57	83.41	43.08	71.24	52.44	57.47	50.00	78.81	50.00	62.20	25.81
DFIL	Task6	98.85	98.90	-	-	-	-	-	-	-	-	-	-	98.85	-
	Task5	27.28	50.49	91.32	90.67	-	-	-	-	-	-	-	-	59.30	71.57
	Task4	67.63	55.59	71.04	60.77	96.17	60.56	-	-	-	-	-	-	78.28	25.75
	Task3	84.93	70.48	76.83	66.73	90.70	57.74	99.05	98.84	-	-	-	-	87.88	11.29
	Task2	83.15	68.99	75.12	61.42	89.69	66.80	81.93	73.83	78.14	77.95	-	-	81.60	13.88
	Task1	39.99	52.74	65.99	66.92	77.30	73.03	90.45	87.09	69.76	70.76	91.50	83.15	72.50	24.01
DS-AL	Task6	99.58	99.41	-	-	-	-	-	-	-	-	-	-	99.58	-
	Task5	99.02	98.21	85.33	89.05	-	-	-	-	-	-	-	-	92.18	0.56
	Task4	98.39	96.91	85.18	87.88	60.57	70.05	-	-	-	-	-	-	81.38	0.67
	Task3	98.07	96.27	85.23	87.82	60.60	69.63	86.01	77.87	-	-	-	-	82.48	0.54
	Task2	97.69	95.48	85.05	87.21	61.49	68.79	85.80	77.64	54.21	51.27	-	-	76.85	0.60
	Task1	97.58	95.27	84.90	86.94	61.58	68.83	85.88	77.55	54.25	50.00	76.37	50.81	76.76	0.51
MMAL	Task6	99.58	99.43	-	-	-	-	-	-	-	-	-	-	99.58	-
	Task5	97.63	95.40	86.46	88.27	-	-	-	-	-	-	-	-	92.05	1.95
	Task4	90.60	80.69	80.93	74.16	90.18	56.60	-	-	-	-	-	-	87.16	7.25
	Task3	89.90	79.83	80.77	73.65	89.78	56.39	86.58	78.64	-	-	-	-	86.76	5.26
	Task2	88.06	76.09	79.70	70.52	90.32	56.67	86.36	78.53	55.85	50.80	-	-	80.06	4.62
	Task1	87.87	75.69	79.42	70.01	90.29	56.66	86.34	78.41	56.52	50.42	78.81	51.52	79.88	3.80
TagFex	Task6	99.28	99.25	-	-	-	-	-	-	-	-	-	-	99.28	-
	Task5	91.78	83.45	84.09	80.39	-	-	-	-	-	-	-	-	87.93	-7.50
	Task4	75.27	50.00	71.51	50.00	96.80	50.00	-	-	-	-	-	-	81.20	-25.89
	Task3	74.65	50.24	71.15	50.24	95.82	51.24	47.05	61.81	-	-	-	-	72.17	-13.89
	Task2	69.08	58.53	73.31	55.69	77.82	61.98	62.82	64.13	71.46	70.30	-	-	70.90	-22.57
	Task1	45.88	50.90	62.61	57.06	70.13	63.23	80.95	69.45	60.98	61.94	87.79	76.25	68.06	-29.92

Table 7: Detailed performance in the modality incremental setting.

Method	Datasets	Video-only ACC	Video-only AUC	Audio-only ACC	Audio-only AUC	Video-Audio ACC	Video-Audio AUC	AAUC ↑	Fk ↓
Finetune	Video-only	93.35	88.38	-	-	-	-	93.35	-
	Audio-only	59.70	67.74	94.86	91.52	-	-	77.28	33.65
	Video-Audio	41.30	58.45	46.76	61.99	54.93	68.34	47.66	50.08
Repaly	Video-only	93.32	88.38	-	-	-	-	93.32	-
	Audio-only	87.97	89.64	94.93	94.34	-	-	91.45	5.35
	Video-Audio	85.58	88.30	88.45	90.45	84.94	88.25	86.33	7.11
iCaRL	Video-only	93.32	88.31	-	-	-	-	93.32	-
	Audio-only	87.63	85.04	92.53	87.76	-	-	90.08	5.69
	Video-Audio	81.43	86.30	88.61	91.29	92.39	93.99	87.48	7.90
LwF	Video-only	93.37	88.44	-	-	-	-	93.37	-
	Audio-only	80.75	78.32	90.32	84.11	-	-	85.54	12.62
	Video-Audio	67.50	77.26	92.30	94.35	83.94	88.76	81.25	12.94
PASS	Video-only	93.33	88.46	-	-	-	-	93.33	-
	Audio-only	58.51	67.03	94.78	91.34	-	-	76.64	34.82
	Video-Audio	34.95	55.44	69.59	78.97	36.35	56.97	46.96	41.79
ACIL	Video-only	94.07	90.62	-	-	-	-	94.07	-
	Audio-only	94.35	90.61	91.88	88.52	-	-	93.12	0
	Video-Audio	77.80	64.79	76.22	64.40	79.17	65.67	77.73	16.11
Fetril	Video-only	93.30	88.31	-	-	-	-	93.30	-
	Audio-only	88.17	87.03	91.77	89.22	-	-	89.97	5.13
	Video-Audio	67.47	53.74	62.80	51.74	76.46	59.19	68.91	27.40
DFIL	Video-only	93.73	90.01	-	-	-	-	93.73	-
	Audio-only	92.13	90.07	94.93	91.79	-	-	93.53	1.60
	Video-Audio	89.27	90.97	90.43	92.42	97.64	97.08	92.45	4.48
DS-AL	Video-only	94.08	90.73	-	-	-	-	94.08	-
	Audio-only	94.22	90.40	92.68	89.25	-	-	93.45	0
	Video-Audio	73.95	61.41	72.48	61.52	79.53	63.35	75.32	20.24
MMAL	Video-only	93.92	90.44	-	-	-	-	93.92	-
	Audio-only	93.80	89.94	92.07	88.22	-	-	92.94	0.12
	Video-Audio	65.70	51.75	59.87	48.46	77.17	58.62	67.58	30.21
TagFex	Video-only	94.12	91.00	-	-	-	-	94.12	-
	Audio-only	91.85	90.00	95.92	92.82	-	-	93.89	2.27
	Video-Audio	82.17	83.37	87.24	87.34	95.46	92.45	88.30	10.32