Scene Background Modelling

Jai Bhatnagar   Hardik Sharma

1 Introduction

Background modeling is a foundational step in computer vision and video analysis, serving as a crucial component for applications such as object detection, motion tracking, video surveillance, and human-computer interaction. The primary goal of background modeling is to distinguish moving objects or dynamic changes from the static or quasi-static background of a scene. This distinction enables the efficient extraction of foreground elements, facilitating further analysis or decision-making processes.

In recent years, the significance of background modeling has grown due to the increasing availability of video data and the demand for real-time processing in various fields. However, the task is not without challenges. Variations in lighting conditions, dynamic background elements (e.g., swaying trees or rippling water), camera motion, and sudden environmental changes make it a complex and evolving problem. Developing robust background modeling techniques capable of adapting to such complexities is essential for improving the reliability and efficiency of downstream applications.

2 Dataset

We have used SceneBackgroundModelling.NET (SBMnet) dataset [1] to evaluate our implemented methods. SBMnet dataset provides a realistic and diverse set of videos. They have been selected to cover a wide range of detection challenges and are representative of typical indoor and outdoor visual data captured today in surveillance, smart environment, and video database scenarios. It provides videos from it's own collection and many other public datasets. Since the dataset is a part of a competition, the authors also provide evaluation scripts which we have used to evaluate our outputs. The dataset provides 8 categories of videos :

- **Basic** category represents a mixture of mild challenges typical of the shadows, Dynamic Background, Camera Jitter and Intermittent Object Motion categories. Some videos have subtle background motion, others have isolated shadows, some have an abandoned object and others have pedestrians that stop for a short while and then move away. These videos are fairly easy, but not trivial, to process, and are provided mainly as reference.
- **Intermittent Motion** category includes videos with scenarios known for causing ghosting artifacts in the detected motion, i.e., objects move, then stop for a short while, after which they start moving again. Some videos include still objects that suddenly start moving, e.g., a parked vehicle driving away, and also abandoned objects. This category is intended for testing how various algorithms adapt to background changes.
- **Clutter** category of videos containing a large number of foreground moving objects occluding a large portion of the background.
- **Jitter** category contains indoor and outdoor videos captured by unstable (e.g., vibrating) cameras. The jitter magnitude varies from one video to another.
- **Illumination Changes**: indoor videos containing strong and mild illumination changes due to a light switch, curtains opening or automatic camera brightness change.
- **Background Motion** category includes scenes with strong (parasitic) background motion: boats on shimmering water, cars passing next to a fountain, or pedestrians, cars and trucks passing in front of a tree shaken by the wind.
- **Very Long**: videos containing more than 3,500 frames.

- **Very Short**: videos containing a limited number of frames (less than 20) with a very low framerate.

3 Methodology

Scene Background Modelling is the task of extracting the background from a given video. Several works have provided many different solutions to this problem. In this project we explore and evaluate various traditional methods to extract the video background without using any modern deep learning based methods.

3.1 Basic Statistical Methods

These methods works on the assumption that foreground objects cover a pixel for a very small number of frames throughout the video. Hence for most part of a video a pixel will have a constant value which is the background and the foreground pixels will be outliers. The methods we have used are :

- **Average** : Since for the majority of the video a pixel will have the background colour and the intensity values can not be too extreme as they are limited to 0-255, therefore averaging out the intensity values for a pixel throughout the video will give us an intensity value close to the background.

- **Median** : Since the foreground pixels will be outliers, the central value should be the background pixel.

- **Mode** : Since the foreground pixels will occur very rarely, the pixel with the highest frequency should be the background pixel.

- **Percentile Average** : In this method we remove a set percentile of outlier values and then take the average. We do this assuming that background pixel may not remain constant but may only show small changes due to environmental factors.

3.2 Frame Difference Segmentation

For this method we perform foreground and background segmentation for each frame using the frame difference method [2]. This includes calculating frame by frame difference in intensity value of each pixel. If the difference is greater than a certain threshold then the pixel is marked as a foreground pixel.

After calculating the background mask for each frame, we take the mean, median or mode of the background pixels over all frames to generate the resulting background image.

3.3 Gaussian Mixture Models

3.4 Optical Flow Modelling

For this method we calculate the optical flow of pixels throughout the video and for each frame we mark the static pixels as background and moving pixels as foreground. We use the inbuilt function in OpenCV library for this. After calculating the background mask for each frame, we take the mean, median or mode of the background pixels over all frames to generate the resulting background image.

3.5 Gaussian Mixture Modelling

Gaussian Mixture Models (GMMs) are a probabilistic approach used to model the pixel intensity distribution of a video frame for background subtraction. The fundamental idea is to represent each pixel as a mixture of Gaussian distributions, where each distribution corresponds to a distinct mode, such as background or foreground. Over time, this enables the system to adapt dynamically to scene changes like lighting variations or background movement.

Mathematical Representation

Let $I_t(x, y)$ denote the intensity of a pixel at position (x, y) in frame t. A GMM models the probability distribution of a pixel as:

$$P(I_t(x,y)) = \sum_{k=1}^{K} w_k \mathcal{N}(I_t(x,y); \mu_k, \sigma_k^2),$$

where:

- K is the number of Gaussian components,

- w_k is the weight of the k-th component, satisfying $\sum_{k=1}^{K} w_k = 1$,

- $\mathcal{N}(I_t(x,y); \mu_k, \sigma_k^2)$ is the Gaussian distribution with mean μ_k and variance σ_k^2:

$$\mathcal{N}(I_t(x,y); \mu_k, \sigma_k^2) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{(I_t(x,y) - \mu_k)^2}{2\sigma_k^2}\right)$$

Background and Foreground Classification

For each incoming pixel value $I_t(x, y)$:

1. **Matching a Component:** The pixel is classified as background if it matches any Gaussian component whose mean μ_k is close to $I_t(x, y)$ within a threshold, here $2.5\sigma_k$.

2. **Updating Parameters:** The matched component's parameters are updated using:

$$\mu_k \leftarrow (1 - \alpha)\mu_k + \alpha I_t(x, y),$$

$$\sigma_k^2 \leftarrow (1 - \alpha)\sigma_k^2 + \alpha(I_t(x, y) - \mu_k)^2,$$

where α is the learning rate.

3. **Handling Unmatched Pixels:** If no component matches, the pixel is classified as foreground, and a new component is initialized or an existing component is replaced with low weight.

The adaptability of GMM allows it to handle dynamic environments, such as waving trees or gradual illumination changes. By weighting and updating Gaussian components, the model maintains a robust and efficient background representation over time.

4 Code Structure

- **evaluation/** - Contains all the code for evaluation as provided by the dataset authors.
- **src/bgModels.py** - Contains the code for all the methods implemented to extract the background.
- **src/main.py** - Contains the pipeline to load a video, call function to extract background and save the results.

5 Experiments

5.1 Metrics

- **AGE**: (Average Gray-level Error). Average of the gray-level absolute difference between GT and the computed background (CB) image. Lower is better.
- **pEPs**: (Percentage of Error Pixels). Percentage of EPs (number of pixels in CB whose value differs from the value of the corresponding pixel in GT by more than a threshold) with respect to the total number of pixels in the image. Lower is better.
- **pCEPS**: (Percentage of Clustered Error Pixels). Percentage of CEPs (number of pixels whose 4-connected neighbors are also error pixels) with respect to the total number of pixels in the image. Lower is better.
- **MSSSIM**: (MultiScale Structural Similarity Index). Estimate of the perceived visual distortion. Higher is better.
- **PSNR**: (Peak Signal to Noise Ratio) Amounts to $10log_10((L - 1)^2/MSE)$ where L is the maximum number of grey levels and MSE is the Mean Squared Error between GT and CB images. Higher is better.
- **CQM**: (Color image Quality Measure). Based on a reversible transformation of the YUV color space and on the PSNR computed in the single YUV bands. It assumes values in db and the higher the CQM value, the better is the background estimate.

5.2 Analysis

5.2.1 Mean

- Strengths: The Average technique performs best in subsets where the background is stable or changes are predictable, such as Background Motion and Very Long sequences. High PSNR and MSSSIM values in these subsets indicate good background reconstruction quality.
- Weaknesses: It struggles in subsets with significant dynamic changes (e.g., Clutter, Illumination Changes). High AGE and low MSSSIM in these subsets highlight the inability to adapt to varying dynamics.
- Overall Trend: Performance degrades as the complexity of the scene increases (e.g., Clutter, Illumination Changes), while stability and predictability in the background enhance accuracy (e.g., Background Motion, Very Long).

5.2.2 Median

- Strengths: The Median technique excels in scenarios with stable or predictable dynamics, such as Background Motion, Basic, and Very Long subsets. Handles illumination changes better than the Average technique, indicating better adaptability to gradual shifts in brightness or contrast.

Category	AGE	pEPs	pCEPs	MSSSIM	PSNR	CQM
backgroundMotion	2.6073	0.0109	0.0032	0.9825	34.5506	33.0699
basic	6.0137	0.0589	0.0364	0.9741	29.0531	30.6344
clutter	21.7032	0.3825	0.2904	0.617	19.6768	21.0804
illuminationChanges	64.4783	0.8371	0.7586	0.8273	10.8803	12.3463
intermittentMotion	10.1341	0.1281	0.103	0.8903	22.5062	23.5667
jitter	7.5481	0.0863	0.0204	0.9212	26.5914	27.725
veryLong	4.2579	0.012	0.0003	0.9887	31.5415	32.2349
veryShort	7.43	0.0862	0.0316	0.9389	26.9296	27.6246

Table 1: Technique-wise Average Metrics (Average)

Category	AGE	pEPs	pCEPs	MSSSIM	PSNR	CQM
backgroundMotion	1.6242	0.0013	0.0001	0.9873	39.637	39.2979
basic	2.5955	0.0109	0.0001	0.9891	35.2965	36.3169
clutter	16.7179	0.2065	0.1551	0.5857	19.4878	20.7494
illuminationChanges	6.755	0.0822	0.0586	0.8765	27.313	28.1198
intermittentMotion	8.3149	0.0743	0.0606	0.8823	21.3906	22.4421
jitter	6.1691	0.0747	0.013	0.9333	27.8006	28.9353
veryLong	3.6466	0.0069	0.0002	0.9933	32.3091	32.9944
veryShort	4.8756	0.0241	0.0003	0.9792	31.8941	32.4441

Table 2: Technique-wise Average Metrics (Median)

- Weaknesses: Struggles in complex or highly dynamic subsets like Clutter, where the presence of overlapping objects reduces its accuracy. Slightly lower performance in shorter sequences compared to longer ones, as it requires sufficient temporal data for optimal convergence.

- Overall Trends: Performs consistently better than Average across most subsets, especially in subsets with minor variations (e.g., Illumination Changes, Intermittent Motion, Jitter). Maintains robustness in filtering transient noise and outliers, making it a reliable choice for stable scenes.

5.2.3 Mode

- Strengths: The Mode technique performs exceptionally well in subsets with repetitive and predictable dynamics, such as Background Motion, Basic, and Very Long. Its ability to focus on the most frequent pixel values makes it robust to transient noise and small variations, such as jitter.

- Weaknesses: Struggles in subsets with rapid or unpredictable variations, such as Illumination Changes and Very Short sequences, where the pixel frequency does not accurately reflect the background. Performs moderately well in Clutter compared to other techniques but is still impacted by overlapping objects.

- Overall Trends: Performs better than Average and Median techniques in subsets with stable dynamics. Shows limitations in subsets requiring adaptability to sudden or complex changes.

5.2.4 Frame Difference Mean

- Strengths: The Frame Difference Average technique excels in subsets with gradual or predictable changes, such as Background Motion, Basic, and Very Long sequences. It performs well in handling small variations caused by Jitter.

- Weaknesses: Subsets with significant pixel variations, such as Illumination Changes and Clutter, present challenges. The averaging approach blurs the distinction between dynamic and static regions, leading to higher errors. Short sequences provide insufficient data for the technique to average out the variations effectively.

- Overall Trends: The performance generally improves with longer sequences or subsets

Category	AGE	pEPs	pCEPs	MSSSIM	PSNR	CQM
backgroundMotion	1.9678	0.0021	0.0	0.9886	37.669	34.5246
basic	2.843	0.015	0.0003	0.9851	33.8383	35.0679
clutter	10.1965	0.0964	0.0236	0.7144	21.6057	22.4532
illuminationChanges	24.706	0.175	0.1066	0.6625	12.4704	13.033
intermittentMotion	7.5035	0.0616	0.0425	0.8938	21.885	22.685
jitter	6.9618	0.08	0.0129	0.9017	24.9013	25.7353
veryLong	3.8777	0.0122	0.0001	0.991	30.8698	31.3568
veryShort	14.8058	0.1948	0.039	0.8373	19.0836	19.6905

Table 3: Technique-wise Average Metrics (Mode)

Category	**AGE**	**pEPs**	**pCEPs**	**MSSSIM**	**PSNR**	**CQM**
backgroundMotion	2.0602	0.0023	0.0004	0.9874	37.1033	34.6793
basic	2.9425	0.0122	0.0003	0.9872	34.6421	35.8188
clutter	17.9059	0.2756	0.2063	0.5861	19.8008	21.1911
illuminationChanges	72.0381	0.8715	0.7994	0.807	10.0057	11.476
intermittentMotion	7.9457	0.0733	0.0598	0.8911	22.1525	23.2233
jitter	6.4956	0.0751	0.0142	0.9323	27.7482	28.8703
veryLong	3.7663	0.0073	0.0001	0.9931	32.3489	33.0024
veryShort	5.5398	0.031	0.0008	0.9695	28.2842	28.9345

Table 4: Frame Difference Average

with stable dynamics. Illumination changes remain a major challenge, highlighting the need for more adaptive techniques to handle such scenarios.

5.2.5 Frame Difference Median

- Strengths:The Frame Difference Median technique excels in subsets with gradual or predictable variations, such as Background Motion, Basic, and Very Long sequences. It handles jitter effectively, smoothing out minor intensity fluctuations.

- Weaknesses:Performance degrades in subsets with high variability or overlap between objects, such as Clutter and Illumination Changes. Median filtering fails to adapt dynamically to such changes. Results are slightly less robust for Very Short sequences due to limited data for accurate filtering.

- Overall Trends:The technique consistently performs better in subsets with longer sequences or stable dynamics, where the median can reliably filter outliers. It struggles in subsets with abrupt intensity changes or significant overlap between foreground and background elements.

5.2.6 Frame Difference Mode

- Strengths:The Frame Difference Mode technique excels in subsets with stable motion or repetitive patterns, such as Background Motion, Basic, and Very Long. It effectively mitigates noise in jittery environments and captures stable dynamics.

- Weaknesses: Performance drops in subsets with high variability, such as Clutter and Illumination Changes, where mode filtering struggles to adapt dynamically. Short sequences, such as Very Short, limit the technique's ability to isolate consistent pixel intensities.

- Overall Trends: The mode consistently performs well in subsets with repetitive patterns or longer sequences, where the most frequent intensity values represent the background. Subsets with dynamic or complex conditions highlight the technique's limitations.

5.2.7 Gaussian Mixture Models

- Strengths: GMM performs well in subsets with stable motion, such as Background Motion and Basic. It handles jitter reasonably well, smoothing out minor pixel fluctuations.

- Weaknesses: The technique struggles with

Category	AGE	pEPs	pCEPs	MSSSIM	PSNR	CQM
backgroundMotion	1.8669	0.0019	0.0002	0.9883	38.2262	35.135
basic	2.5564	0.0102	0.0001	0.9892	35.3096	36.4053
clutter	16.756	0.1758	0.1291	0.5657	18.6434	19.9459
illuminationChanges	7.736	0.1136	0.0495	0.8242	25.6921	24.7341
intermittentMotion	8.4891	0.0714	0.0575	0.8797	21.0824	22.1433
jitter	6.1964	0.0722	0.011	0.9305	27.4789	28.6386
veryLong	3.6427	0.0066	0.0002	0.9933	32.2573	32.9458
veryShort	5.1337	0.0283	0.0002	0.9769	31.0388	31.6157

Table 5: Frame Difference Median

Category	AGE	pEPs	pCEPs	MSSSIM	PSNR	CQM
backgroundMotion	1.9328	0.0017	0.0001	0.9894	38.0297	34.9246
basic	2.7046	0.0116	0.0001	0.9874	34.4567	35.5359
clutter	15.1302	0.1421	0.0737	0.6033	18.9254	19.8772
illuminationChanges	23.099	0.1638	0.0944	0.6459	13.1199	13.6372
intermittentMotion	7.9421	0.0657	0.0489	0.8816	21.2995	22.1689
jitter	6.7644	0.077	0.0109	0.9084	25.4251	26.4379
veryLong	3.6996	0.0089	0.0002	0.993	31.9464	32.5938
veryShort	9.1132	0.1084	0.0085	0.9329	24.4381	25.0606

Table 6: Frame Difference Mode

Category	AGE	pEPs	pCEPs	MSSSIM	PSNR	CQM
backgroundMotion	7.7353	0.106	0.0183	0.9357	25.2753	25.5579
basic	7.0618	0.0976	0.015	0.9285	25.6889	27.1197
clutter	18.3839	0.3038	0.2094	0.5853	20.2435	21.5976
illuminationChanges	62.1728	0.8314	0.7229	0.7971	11.1469	12.5902
intermittentMotion	10.9972	0.1349	0.0788	0.8581	21.5545	22.6815
jitter	6.9973	0.092	0.0227	0.9204	27.3601	28.4935
veryLong	16.1468	0.2375	0.0512	0.8442	19.0827	20.4123
veryShort	15.2749	0.2481	0.0566	0.8707	20.7566	21.8196

Table 7: Performance metrics for GMM.

high variability, as seen in Clutter and Illumination Changes. Long and short sequences (Very Long and Very Short) highlight the difficulty of adapting to dynamic or limited data.

- Overall Trends: GMM adapts well to subtle and consistent motion patterns, but its reliance on Gaussian distributions limits its effectiveness in handling dynamic, cluttered, or abrupt changes.

5.2.8 Optical Flow Average

- Strengths:

  Background Motion, Very Long, and Basic subsets highlight the strong performance of Optical Flow Average in scenarios with consistent or predictable motion. Handles jitter well, ensuring smooth and accurate representation in minor variations.

- Weaknesses:

  Struggles in highly dynamic environments such as Clutter and Illumination Changes, with significant errors and reduced visual quality. Elevated AGE in subsets with complex motion patterns, such as Intermittent Motion and Very Short.

- Overall Trends:

  Optical Flow Average thrives in scenes with gradual or consistent motion changes, leveraging its ability to track optical flow vectors effectively. It falters in subsets requiring adaptation to abrupt or high-intensity variations, such as illumination changes or complex clutter.

5.2.9 Optical Flow Median

- Strengths: Outstanding performance in Background Motion, Basic, and Very Long subsets due to its capability to handle gradual changes effectively. High MSSSIM and PSNR across subsets indicate strong visual quality retention.

- Weaknesses: Struggles in dynamic scenarios such as Clutter and Illumination Changes, with elevated AGE and error rates. Performance declines in complex motion patterns, likely due to the smoothing effect of median filtering.

- Overall Trends: Optical Flow Median performs exceptionally in stable, predictable motion scenarios, offering high-quality results with minimal artifacts. It is less effective in subsets requiring rapid adaptation to abrupt changes or high dynamic complexity.

5.2.10 Overall Conclusion

In analyzing the performance of various background modeling techniques across the dataset, it becomes evident that methods such as Median, Frame Difference Median, and Optical Flow Median consistently outperform the others. These techniques excel due to their ability to minimize errors (such as AGE, pEPs, and pCEPs) while maintaining high perceptual quality in terms of MSS-SIM, PSNR, and CQM. The key strength of these methods lies in their ability to effectively handle dynamic changes in the scene, reducing distortions and preserving background accuracy.

Median-based methods, in particular, achieve a balance of low error rates and high image quality. This trend is particularly noticeable in techniques like Frame Difference Median and Optical Flow Median, which show robust performance across all metrics. These methods are especially adept at maintaining image quality under varying conditions, likely due to their ability to effectively process outliers and noise.

On the other hand, techniques like GMM and Pooled GMM Frame Diff Avg show a noticeable decline in performance. These methods suffer from high AGE and pEPs, leading to poor error rates and reduced image quality. The underlying issue seems to be the difficulty these techniques face in managing sudden scene changes or background complexities, which results in larger discrepancies in pixel-level changes.

Methods that rely on average or most frequent approaches, like Frame Difference Average and Mode, provide moderate performance but fail to match the quality of median-based methods. These techniques often struggle with noise or rapid changes, resulting in lower quality visual outputs.

In conclusion, the superior performance of Median-based methods is attributed to their robustness against errors and superior ability to maintain visual integrity, making them the most reliable choice for background modeling tasks across diverse video datasets.

Category	AGE	pEPs	pCEPs	MSSSIM	PSNR	CQM
backgroundMotion	2.8241	0.0117	0.0034	0.9823	29.2643	29.0997
basic	6.2978	0.0668	0.0415	0.973	28.7792	30.3429
clutter	21.6902	0.3827	0.29	0.6181	19.6864	21.0955
illuminationChanges	64.6061	0.8371	0.7583	0.8264	10.8628	12.3291
intermittentMotion	10.1341	0.1281	0.1031	0.8904	22.5067	23.5678
jitter	7.5488	0.0863	0.0204	0.9211	26.5871	27.7205
veryLong	4.2563	0.012	0.0003	0.9887	31.5446	32.2367
veryShort	7.9347	0.1003	0.0411	0.9319	26.2889	26.9898

Table 8: Performance metrics for Optical Flow Average.

Category	AGE	pEPs	pCEPs	MSSSIM	PSNR	CQM
backgroundMotion	4.7067	0.0514	0.002	0.9625	28.0372	28.0745
basic	5.316	0.0561	0.0054	0.9568	27.6353	29.0385
clutter	19.9818	0.3135	0.234	0.5895	19.3579	20.7557
illuminationChanges	72.7804	0.8749	0.8067	0.8004	9.9194	11.3909
intermittentMotion	9.2589	0.1082	0.0718	0.8751	21.6228	22.7445
jitter	6.5953	0.0813	0.0182	0.9318	27.7021	28.8244
veryLong	10.0587	0.1249	0.0064	0.8849	21.8319	23.0251
veryShort	10.3353	0.1411	0.02	0.9298	23.1965	24.1521

Table 9: Performance metrics for Pooled GMM-Frame Difference Average.

Category	AGE	pEPs	pCEPs	MSSSIM	PSNR	CQM
backgroundMotion	4.7067	0.0514	0.002	0.9625	28.0372	28.0745
basic	5.316	0.0561	0.0054	0.9568	27.6353	29.0385
clutter	19.9818	0.3135	0.234	0.5895	19.3579	20.7557
illuminationChanges	72.7804	0.8749	0.8067	0.8004	9.9194	11.3909
intermittentMotion	9.2589	0.1082	0.0718	0.8751	21.6228	22.7445
jitter	6.5953	0.0813	0.0182	0.9318	27.7021	28.8244
veryLong	10.0587	0.1249	0.0064	0.8849	21.8319	23.0251
veryShort	10.3353	0.1411	0.02	0.9298	23.1965	24.1521

Table 10: Performance metrics for Pooled GMM-Frame Difference Average.

References

[1] Pierre-Marc Jodoin, Lucia Maddalena, Alfredo Petrosino, and Yi Wang. Extensive benchmark and survey of modeling methods for scene background initialization. *IEEE Transactions on Image Processing*, 26(11):5244–5256, 2017.

[2] Meijin Li, Ying Zhu, and Jiandeng Huang. Video background extraction based on improved mode algorithm. In *2009 Third International Conference on Genetic and Evolutionary Computing*, pages 331–334, 2009.

Category	**AGE**	**pEPs**	**pCEPs**	**MSSSIM**	**PSNR**	**CQM**
backgroundMotion	3.5218	0.0297	0.0006	0.976	30.2218	29.9698
basic	4.4157	0.0418	0.0029	0.969	28.8302	30.0455
clutter	20.3134	0.3009	0.2185	0.5695	18.9739	20.2709
illuminationChanges	6.7835	0.0805	0.0526	0.8866	27.2444	28.0604
intermittentMotion	9.1577	0.102	0.071	0.8709	21.3154	22.3998
jitter	6.3708	0.0798	0.0168	0.9326	27.7716	28.8932
veryLong	7.8187	0.0807	0.0036	0.9327	23.5661	24.3954
veryShort	9.3091	0.121	0.0103	0.9424	24.1358	24.9539

Table 11: Performance metrics for Pooled GMM-Frame Difference Average-Median.

Method Metric	AGE	pEPs	pCEPs	MSSSIM	PSNR	CQM
Average	15.5236	0.2003	0.1555	0.8925	25.2168	25.9969
Median	**6.3540**	**0.0601**	0.0361	**0.9034**	**29.3876**	**30.1175**
Mode	9.1265	0.0798	**0.028**	0.8717	25.3002	25.4466
FrameDifferenceAverage	14.8395	0.1686	0.1353	0.8942	26.5103	27.1092
FrameDifferenceMedian	6.5568	0.0603	0.0312	0.8932	28.7085	28.8898
FrameDifferenceMostFrequent	8.8200	0.0729	0.0294	0.8670	25.9642	26.1803
GMM	18.0963	0.2564	0.1469	0.8425	21.3886	22.5340
OpticalFlowAverage	15.6636	0.2031	0.1573	0.8915	24.4404	25.3853
OpticalFlowMedian	7.2259	0.0810	0.0477	0.8814	28.7460	29.2815
PercentileAverage	14.5404	0.1690	0.1373	0.8949	27.3137	28.1606
PooledGMMFrameDiffAvg	17.3791	0.2189	0.1456	0.8664	22.4129	23.5007
PooledGMMFrameDiffAvgMedian	8.4613	0.1046	0.0470	0.8850	25.2574	26.1236

Table 12: Overall Average Metrics