Video Blur Detection

Anurodh Jain, Pragya Roy, Ronak Dedhiya

***Abstract*— In ubiquitous multimedia area, the number of digital videos increases dramatically with various qualities in video frames. Artifacts such as blur may commonly exist in videos which will disturb compression and retrieval applications. Blur is one of the conventional image quality degradations which is caused by various factors like limited contrast, inappropriate exposure time and improper device handling. For video blur detection, blur measure is calculated for each of the video frame followed by computing statistical measure which forms our input feature sets. We examine different combinations of features and train various machine learning algorithms with hyperparameter tuning. We perform qualitative comparison of the available state-of-the-art methods. Performance is measured and compared in terms of precision, recall, f-measure, accuracy, and execution time demonstrating the effectiveness of each of the techniques. Based on the computed score, Laplacian performs better than other techniques and a combination of different blur measure further improves the classification results.**

I. INTRODUCTION

With the exponential growth of video capturing devices, studies on characterization and detection for blur regions from digital image have become one of the most important research branches in recent years [1-2]. In addition to the use as a part of de-blurring process, automatic detection, and classification of the blurred regions from digital image are very functional to understand the image information and useful for evaluating image quality for further enhancement processes [4]. Image blurriness can be categorized into motion blur and defocus blur. The motion blur can occur due to two potential reasons, when moving objects are captured and when camera is in motion either intentionally or unintentionally. Whereas the defocus blur usually occurs to highlight the focus and out-of-focus regions in the image [5]. An image contains useful information that can be used in various computer vision and image processing applications, i.e., background tracing, text retrieval, image retrieval, person authentication, etc. However, blur affects the contrast and sharpness details of the image that made the retrieval of information challenging.

Image based blur detections techniques are found to be very powerful and readily can be extended to video data. These techniques assume measurement of blurriness at each frame, without considering the motion information or motion induced blurriness as shown in Figure 1.

The aim of the study is to understand available state of the arts blur measure techniques. Implement the techniques and employ it to the video frame. Our main contribution lies in using statistical measure on the vector of blur measure values obtained from video and training various machine learning algorithm for classifying blurry vs non-blurry video. This paper focuses on the different blur detection techniques and aims to compare the performance of each one in terms of accuracy rate, precision, recall, f1 score and execution time. We identify the best set of features along with high performing machine learning algorithm with hyperparameter tuning for each of the blur detection methodology. We also extend the study by combining feature values from this blur detection methodology and training a machine learning algorithm which is found to be outperforming all individual methods.

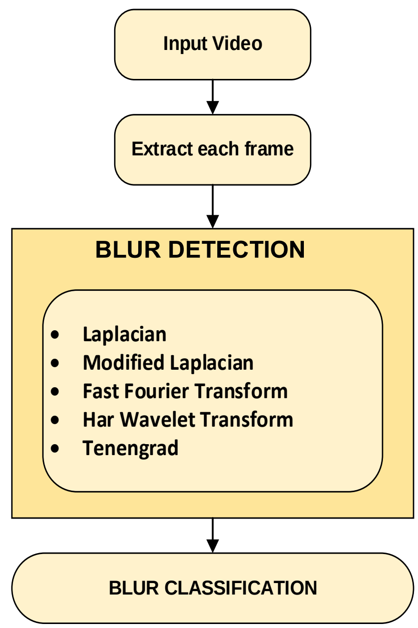

Figure 1 Complete framework

II. DATASET DESCRIPTION

The dataset provided for this project consists of 25 reference videos of 10 seconds duration in HD and associated distorted versions [13]. These videos have been artificially degraded by means of distortion generation algorithms. The frame rate is 30fps. The dataset is divided for training and testing the model. 18 video scenarios (for each scenario there is one original video along with 4 distorted videos where the distortion level varies as – just noticeable, visible but not annoying, annoying, very annoying) are used for training the model and 7 video scenarios are used for testing the model.

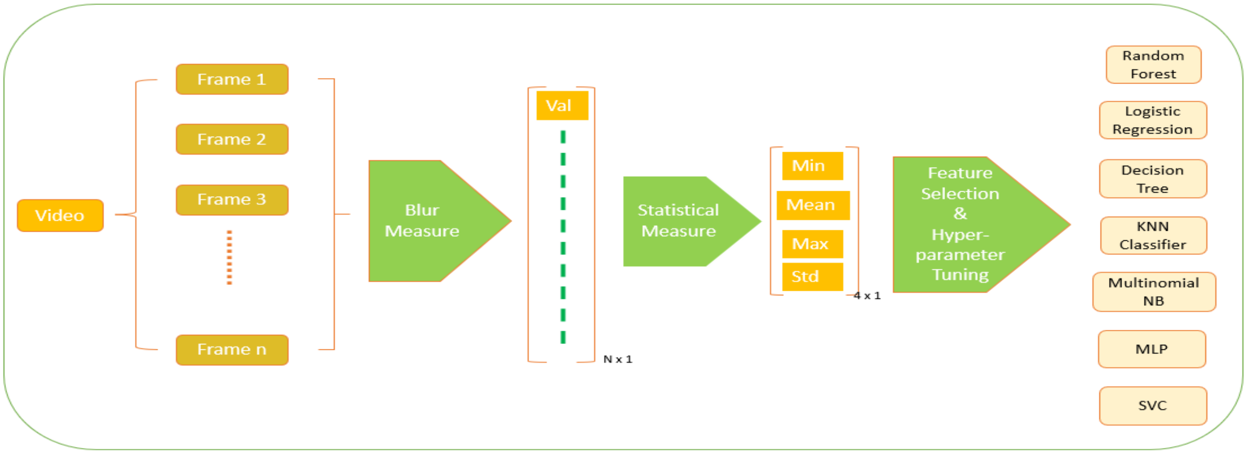

Figure 2: End to End Video Blur Detection Pipeline

III. METHODOLOGY

The paper describes in detail the pipeline of end-to-end video blur detection methodology as shown in Figure 2. The final machine learning models is trained to classify blurry vs non-blurry videos. Each of the modules are discussed below in detail:

A. Video to Frames

The videos are extracted at 30fps, and each video is of 10 seconds. So, the total number of frames in a video are 300.

B. Blur Detection

The Blur images retain some of the features but those are spread out to a wider area hindering the data of the neighboring pixels. When a smooth image's gradient curve is compared to blurred image's gradient curve, the gradient for the blurred images moves faster in comparison to that of the smooth image, thus, making the second derivative larger for the blurred image. As the rate of change of the rate of change of the pixel values in an image is very high, there is blurriness in the image.

Below are the Blur detection techniques:

1) Laplacian:

The Laplace operator of discrete function is obtained by taking difference of the second derivative of Laplace operator in X and Y directions. The Laplacian highlights regions of an image containing rapid intensity changes. The assumption here is that if an image contains high variance, then there is a wide spread of responses both edge-like and non-edge like representative of a normal in-focus image [3]. But if there is very low variance then there is small spread of responses, indicating there are very little edges in the image. The more an image is blurred the less edges there are.

2) Modified Laplacian:

The modified Laplacian is developed to compute local measures of the quality of image focus. By getting the absolute values of the second derivatives in x and y directions. The Modified Laplacian method explores the Laplacian operator in a different fashion. Instead of looking at the variance it looks at the absolute values of the filtered image. The interpretation is that, in sharper images on average large values (both negative and positive) of its Laplacian.

3) Fast Fourier Transform:

The Fast Fourier Transform is a convenient mathematical algorithm for computing the Discrete Fourier Transform. It is used to decompose an image into its sine and cosine components. The output of the transformation represents the image in the Fourier or frequency domain, while the input image is the spatial domain equivalent. In the Fourier domain image each point represents a particular frequency contained in the spatial domain image. The number of frequencies corresponds to the number of pixels in the spatial domain image. This method calculates the frequencies in the image at different points and based on the set level of frequencies it decides whether the image is blurred or sharp. When there is a low amount of frequency based on the set level of frequencies then it declares that the image is blurred otherwise, if the computed frequencies is high then the image is sharp.

4) Haar Wavelet Transform:

A Wavelet is a wave-like oscillation that is localized in time. The algorithm classifies an image as blurred or sharp by splitting the image into N x N tiles, applying several iterations of the 2D HWT to each tile, and grouping horizontally, vertically, and diagonally connected tiles with pronounced changes into tile clusters [7]. Images with large tile clusters are classified as sharp. Images with small tile clusters are classified blurred.

5) Tenengrad:

In image processing, an edge is a region of quick color (or brightness in grayscale images) change. The quicker the change sharper the edge. Neighboring pixels are compared to find the edge [6]. Change in pixel can be summarized as:

- Darker pixels mean negative change.
- Brighter pixels mean positive change
- Grey pixels mean neighbors look roughly alike.

With both the horizontal and vertical gradient components, the region where there is maximum brightness change can be identified. Both the horizontal and vertical when combined, generates the magnitude (measure of change of brightness) for a given pixel, regardless of direction. This is the concept behind Sobel filter. The Tenengrad method, relies on the magnitude of Sobel filter. The Tenengrad builds on the fact that, on average, sharper images will produce larger gradient magnitudes when compared with blurry images.

C. Blur Measure

The blur detection techniques give us the blur measure (amount of blur) for each of the video frame. For every video, we get a N x 1 vector (i.e.,300 x 1) vector.

D. Statistical Measures

Since the blurriness in the video is more or less uniformly distributed, instead of using vector of 300 values, we can use statistical measures, replace the N x 1 vector with min, max, mean and standard deviation which represent the distribution of the vector and can be used readily for classification.

E. Feature Selection & Hyper parameter Tuning

We primarily have 4 features and can further select feature among them thus we tried all the possible combination of features and trained the various classification machine learning algorithms with Hyper parameter tuning using Sklearn's GridsearchCV.

F. Machine Learning algorithm

We tried various types of Machine learning models including Logistic regression, Decision trees, random forest, KNN Classifier, Multinomial Naïve Byes, Multilayer Perceptron, and SVC[8-12].

IV. EVALUATION MEASURES

Accuracy, Recall, Precision, F1 Scores and execution time are metrics that are used to evaluate the performance of a model.

Recall is the ability of a model to find all the relevant cases within a data set. Recall can be defined as the number of true positives divided by the number of true positives plus the number of false negatives.

Precision is the ability of a classification model to identify only the relevant data points. Precision can be defined as the number of true positives divided by the number of true positives plus the number of false positives.
As we increase precision, we decrease recall and vice-versa.

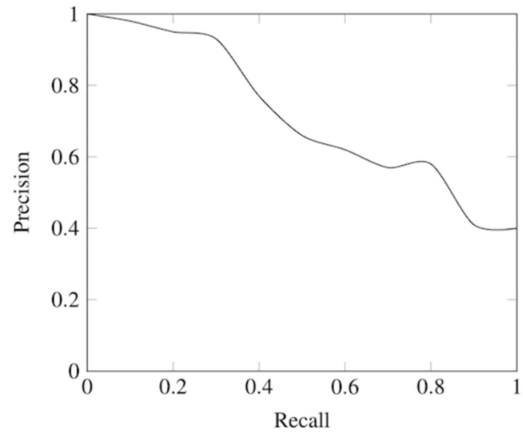

Figure 3 precision vs recall

The **F1 score** is the harmonic mean of precision and recall.
To create a classification model with the optimal balance of recall and precision, then F1 score is maximized.

Accuracy measures the fraction of correct predictions. It is the ratio of correct predictions to total predictions made. It is a great measure but only when datasets are symmetric i.e., where values of false positive and false negatives are almost same.

Execution time can be defined as the measure of time from program initiation at presentation of inputs to termination at the delivery of the outputs.

V. EXPERIMENTS & RESULTS

We used available state of the art blur detection techniques as mentioned in section II. Each of these techniques has four features (min, max, mean, and standard deviation). We tried different combinations of these features. We used various machine learning models and trained them with these features. Then we tuned the hyperparameters for each of these models and based on that, used the best model for each of the blur detection techniques to get the results. We did a qualitative comparison of the machine learning algorithms for all the blur detection techniques with different combination of features. We also tried with different combinations of blur detection techniques to get the best results.

Following are the results obtained with Laplacian blur measure:

Classifer	Features Used	Parameters	Accuracy	Recall	Precis ion	F1S core
Random Forest	min	Depth=2, Estimator=10	0.767	0.964	0.9	0.93 1
Logistic Regressi on	min	C = 0.1, l2, balanced	0.767	0.964	0.9	0.93 1

Decision Tree	min, mean, std	Balanced, depth=4, min split=3	0.821	0.928	0.928	0.928
KNN Classifier	min,mean,max	N_neigbours = 4	0.83	0.964	0.93	0.947
Multinomial NB	min,max	alpha = 1	0.696	0.964	0.87	0.915
MLP	mean, std	alpha=5e-3, hidden = (10,10,10)	0.839	0.964	0.93	0.947
SVC(Poly)	mean, std	C=10,balanced, degree = 2	0.839	0.964	0.93	0.947

All models give satisfying results, results are obtained from single feature as well.

Following are the results obtained with modified Laplacian measure:

Classifer	Features Used	Parameters	Accuracy	Recall	Precision	F1Score
Random Forest	min	None, depth=2, estimator=30	0.5	1	0.8	0.889
Logistic Regression	min,mean,max,std	C=0.01, Balanced, l2	0.607	0.5	0.875	0.636
Decision Tree	min	Balanced, depth=2, split=2	0.5	1	0.8	0.889
KNN	min	Neighbours = 5	0.5	1	0.8	0.889
Multinomial NB	min	alpha = 10	0.5	1	0.8	0.889
MLP	min	adam, layer size=10, 0.001	0.5	1	0.8	0.889
SVC(Poly)	min	C=1,None, degree=3	0.5	1	0.8	0.889

Classifiers accuracy is less for all the different machine learning algorithm for modified Laplacian measure. It is less discriminatory compared to Laplacian.

Following are the results obtained with FFT:

Classifer	Features Used	Parameters	Accuracy	Recall	Precision	F1Score
Random Forest	min,max	Depth=4, Estimator=30,balanced	0.767	0.964	0.9	0.931
Logistic Regression	min,mean	C= 0.01,Balanced, l2	0.75	0.928	0.896	0.912
Decision Tree	max,std	Balanced, depth=3, min split=2	0.73	0.89	0.89	0.89
KNN	max	Neighbours = 2	0.803	0.89	0.925	0.91
Multinomial NB	std	alpha = 1	0.5	1	0.8	0.89
MLP	min,mean,max,std	alpha=0.05, hidden layer=10	0.642	1	0.848	0.918
SVC(Poly)	mean	balanced, degree = 2, C = 1	0.75	0.928	0.896	0.912

Following results are obtained with HWT:

Classifer	Features Used	Parameters	Accuracy	Recall	Precision	F1Score
Random Forest	min,max	Depth=2, Estimator=30,None	0.767	0.964	0.9	0.931
Logistic Regression	min	C=1,None, l2	0.75	0.928	0.896	0.912
Decision Tree	max	Balanced, depth=4, min split=4	0.803	0.898	0.925	0.9
KNN	max or mean_std	neighBour = 2	0.75	0.928	0.896	0.912
Multinomial NB	max	alpha = 10	0.5	1	0.8	0.889
MLP	min,max	adam,(10,10), 5e-3	0.767	0.964	0.9	0.931
SVC(poly)	mean,std	Balanced, degree = 3, C = 0.01	0.857	0.857	0.96	0.905

And finally following results are obtained with Tenengrad:

Classifer	Features Used	Parameters	Accuracy	Recall	Precision	F1Score
Random Forest	mean,std	balanced, depth=2, estimators = 30	0.75	0.928	0.896	0.912
Logistic Regression	mean,std	balanced, l2, C = 10	0.73	0.89	0.892	0.892
Decision Tree	max	Depth=3, min split=3, None	0.696	0.964	0.87	0.915
KNN	max	neighbours = 3	0.696	0.964	0.87	0.915
Multinomial NB	mean,max	alpha = 1	0.57	1	0.823	0.903
MLP	max,std	1e-3, (10,10,10), adam	0.625	0.964	0.843	0.9
SVC(Poly)	max	Balanced, C = 0.01, Degree = 2	0.732	0.892	0.892	0.892

We also do final comparison among the methods and combination of theses techniques whose results are as discussed below:

Blur Detection Technique	Classifer	Features Used	Parameters	Accuracy	Recall	Precision	F1Score	Execution Time (sec)
FFT	Random Forest	min,max	Depth=4, Estimator=30, balanced	0.767	0.964	0.9	0.931	95
Laplacian	SVC(Poly)	std	C=10,balanced, degree = 2	0.839	0.964	0.93	0.947	17
Modified Laplacian	Random Forest	min	None, depth=2, estimator=30	0.5	1	0.8	0.889	19
HWT	MLP	min,max	adam,(10,10), 5e-3	0.767	0.964	0.9	0.931	76
Tenengrad	KNN	max	neighbours = 3	0.696	0.964	0.87	0.915	68
FFT + laplacian + modified laplacian	Randome Forest	min	depth =2, estimators = 10, none	0.857	1	0.933	0.9657	131
Laplacian + modified laplacian	Random Forest	Laplacian(max, std), Modified Laplacian(mean)	Depth = 4, estimator = 30	0.857	1	0.933	0.9657	36

Here,we see combination of features outperform the individual methods.

VI. CONCLUSION

Among all the blur detection techniques, Laplacian provided the best f1-score. When combining different techniques, we observed that combinations of Laplacian + Modified Laplacian and combinations of FFT + Laplacian + Modified Laplacian performed the best with f1-score of 0.9657, however the overall execution time of the video processing was significantly less if we do not have FFT in the combination.

VII. REFERENCES

[1] Su, B., Lu, S., Tan, C. L., 2011. Blurred image region detection and classification, in MM'11 - Proceedings of the 2011 ACM Multimedia Conference and CoLocated Workshops.
[2] Yang, F., Lin, H. J., Chuang, H., 2017. Image deblurring, 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation(SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), San Francisco, CA.
[3] Bansal, R., Raj, G., Choudhury T., 2017. Blur image detection using Laplacian operator and Open-CV, Proc. 5th Int. Conf. Syst. Model. Adv. Res. Trends, SMART 2016, pp. 63–67.
[4] Huang, R., Fan, M., Xing, Y., Zou, Y., 2019. Image Blur Classification and Unintentional Blur Removal, IEEE Access.
[5] Koik, B. T., Ibrahim, H., 2014. A literature survey on blur detection algorithms for digital imaging, Proc. - 1st Int. Conf. Artif. Intell. Model. Simulation, AIMS 2013, pp. 272–277.
[6] Gao, S., Han, M., Cheng, X., 2018. The fast iris image clarity evaluation based on Tenengrad and ROI selection.
[7] Tong, H., Li, M., Zhang, H., Zhang, C., 2004. Blur detection for digital images using wavelet transform, in 2004 IEEE International Conference on Multimedia and Expo (ICME).
[8] (2011) Logistic Regression. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-30164-8_493
[9] Fürnkranz J. (2011) Decision Tree. In: Sammut C., Webb G.I. (eds) Encyclopedia of Machine Learning. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-30164-8_204
[10] Liu, Y., Wang, Y., Zhang, J. (2012). New Machine Learning Algorithm: Random Forest. In: Liu, B., Ma, M., Chang, J. (eds) Information Computing and Applications. ICICA 2012. Lecture Notes in Computer Science, vol 7473. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-34062-8_32
[11] Lishan Wang 2019 IOP Conf. Ser.: Mater. Sci. Eng. 677 052038
[12] Kibriya, A.M., Frank, E., Pfahringer, B., Holmes, G. (2004). Multinomial Naive Bayes for Text Categorization Revisited. In: Webb, G.I., Yu, X. (eds) AI 2004: Advances in Artificial Intelligence. AI 2004. Lecture Notes in Computer Science(), vol 3339. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-30549-1_43
[13] Video blur dataset [accessed 16/04/2022], https://vsquad2022.aliqureshi.info/index.html, https://www.l2ti.univ-paris13.fr/VSQuad/
[14] Pagaduan, Roxanne A., Ma Christina R. Aragon, and Ruji P. Medina. "iblurdetect: Image blur detection techniques assessment and evaluation study." In Proceedings of the International Conference on Culture Heritage, Education, Sustainable Tourism, and Innovation Technologies-CESIT, pp. 286-291. 2021.