ISSN No:-2456-2165

Optimization of Facial Recognition for Surveillance and Security: A Mathematical Framework based on the Minimum Cross Entropy Theory

Tamara Regina[1], Christopher Matthew[2]
[1,2]International Business Engineering Program, Petra Christian University, Indonesia

Abstract:- The development of effective surveillance and security systems is crucial for public safety and for preventing criminal activities. However, lack of proper infrastructure and technological advancements has hindered the progress of surveillance and security systems in many parts of the world. One significant challenge in this regard is the difficulty in suspect identification from manually inspecting several footags. In this paper, we propose the optimization of facial recognition systems, commonly used in surveillance and safety, through several image processing techniques such as Eigen Faces, Fisher Faces, and Facial Coordinates based on the minimum cross entropy theory. Our approach involves several features incorporating facial recognition that can be analyzed to identify potential suspects. By the end of this paper, we found an interesting result that only one can accurately predict facial recognition from video footage given the training image.

I. INTRODUCTION

In an era characterized by technological advances and groundbreaking discoveries, facial recognition has emerged as a state-of-the-art technology with a high potential to revolutionize various aspects of our lives, including in the sector of surveillance and security. Facial recognition is a biometric technology that adopts sophisticated algorithms for person identification based on their unique facial features. It leverages computer vision and machine learning techniques to analyze and match facial characteristics, mainly structural features, shape, and texture of a person's face. By comparing captured facial images or video frames against a database of images, this technology can determine the identity of an individual, which is particularly useful in surveillance and security for suspect identification (Muftić, 2021). From several years of development, significant advances have been made in the design of classifiers for successful face recognition. Among appearance-based holistic approaches, Eigenfaces (Turk & Pentland, 1991) and Fisherfaces (Anggo et.al, 2018) have proved to be effective in experiments with large databases. With traditional methods, face recognition can create matching errors due to invariance factors from pose, hairstyle, different physical conditions, etc. Therefore, face recognition based on facial geometric landmarks was developed. A recent study on facial recognition based on facial landmarks was conducted and proved to solve these problems. In analyzing these features for facial recognition system, optimization problem is needed to ensure greater accuracy in identification of a person. Whilst these features are widely used in facial recognition technology, we proposed a new method that utilizes cross entropy as our minimization problem as a parameter to calculate loss between two distributions, mainly expected and observed distributions. Cross entropy is a measure of how well the observed distribution matches the expected distribution (Brownlee, 2020). This means that the optimization algorithm is trying to find the probability distribution of the observed features that are as close as possible to the expected distribution of features. In our experiment, the expected distribution will be the distributions of our images from the database, where the observed distribution is the distribution of images from our video footage, in the context of suspect identification in surveillance and security. In addition, to calculate dissimilarity of two distributions in each variable, we will compute our cross entropy based on the Mahalanobis distance, which are proven to be a crucial aspect in characteristics differentiation (Zhang et al., 2020). Promising results are shown at the end of this paper. The paper is organized as follows : Introduction is in Section I and Section II describes literature review. Section III is an introduction to our methods in our optimization problem. Experimental results are presented in Section IV and Section V for sensitivity analysis. Section VI describes our conclusion and future work.

II. LITERATURE REVIEW

Surveillance and security services have been evolving around for centuries, with the earliest recorded security measures back in the 625 BC (Gentile, 2023). However, the rapid growth of surveillance and security services integrated with sophisticated advancements of technology have provided many protection for individuals, businesses, governments, and ultimately the whole society. In the 20th century, security services evolved to include technology such as CCTV which improved the effectiveness of many parties including officers, police departments, government, enabling them to monitor and respond to any security threats efficiently.

The impact of technology on the surveillance and security systems cannot be overstated. The proliferation of digital technology has led to the development of sophisticated security systems that can detect and deter threats in real time. These technological advancements improved the effectiveness of security officers, enabling them to monitor and respond to threats efficiently. These systems have transformed the role of security employees from passive observers to proactive responders.

Without doubt, surveillance cameras have become an essential tool for public safety and security (Insider, 2021). With the increasing crime rate, these cameras are an efficient way to keep track of suspicious activities and identify criminals. However, even until today, typical investigators has to sift through masses of arrest records, or even ask outsiders to help identify people captured in footage. In reality, many agencies lack tools to effectively compare individuals caught in thousands of footages against photos of known offenders and previously arrested individuals. Presently, numerous law enforcement agencies rely on arduous and labor-intensive manual procedures to discern individuals depicted in videos. Consequently, due to resource constraints, these agencies often prioritize investigations of grave or high-profile offenses, inadvertently leaving a significant number of cases unresolved. A study of police criminal investigation practices has shown that much of the typical investigator's time is devoted to manually screening information and that eyewitness testimony sometimes incur identification errors due to memory 'contamination' (Hildreth & Rueda, n.d.) With smart access to advanced surveillance systems and unprecedented volumes of video footage, the manual techniques used to sift through it all do not efficiently scale at the same rate. Thus, an automated system that can quickly and accurately detect the presence of a criminal in surveillance footage would be a revolutionary step towards improving public safety.

Facial recognition technology possesses the capability to swiftly and accurately ascertain individuals in real-time, thus enabling the timely notification of agencies regarding potential security risks. Considering the vitality of obtaining high accurate results in face recognition and suspect identification, the presence of a robust and effective software is highly demanded. The use of minimum cross entropy is highly useful in the case of increasing accuracy of face recognition systems by minimizing the discrepancy between predicted and actual face features. This statistical method, will help to optimize the model's parameters to better align with ground truth data, enhancing its ability to recognize and differentiate individuals accurately.

III. RESEARCH METHODOLOGY

The objective of the optimization problem is to minimize the cross-entropy loss between the expected distribution and the observed distribution of the features extracted from the images. Cross-entropy is a measure of the difference between two probability distributions. In the context of image processing in surveillance and security, the expected distribution represents the distribution of features we observe in the given image, while the observed distribution represents the distribution that we observed in video frames.

The cross-entropy loss is a measure of how well the observed distribution matches the expected distribution. The objective of the optimization problem is to minimize this cross-entropy loss. This means that the optimization algorithm is trying to find the probability distribution of the observed features that is as close as possible to the expected distribution of features. By minimizing the cross-entropy loss, the algorithm is effectively minimizing the difference between the two distributions and improving the accuracy of the image processing system. The cross-entropy loss will be zero if the observed and expected distributions are identical. In our context, 0 cross entropy means correct identification of suspects from the footage given the image of the suspect in the database. However, if the observed distribution deviates from the expected distribution, the cross-entropy loss will be positive. The larger the deviation between the two distributions, the larger the value of the cross-entropy loss.

The mathematical formula for the objective function is given by:

$$\text{Minimize } H(P, Q) = -\sum_i P_i \log(Q_i)$$

where Pi is the probability of feature i according to the expected distribution P, and Qi is the probability of feature i according to the observed distribution Q

Subject to:

- The sum of the probabilities of the features should be equal to one:
- $\text{sum}(Q_i) = 1$
- The probabilities of the features should be non-negative:
- $Q_i >= 0$

In our case expected distribution is the probability distribution of the features we expect to see in the image, whilst the observed distribution is the probability distribution of the features we actually observe in the video stream.

In computing the minimum cross entropies, several features are trialed to create the minimization of entropy (least) and see which several combinations of features best create minimum loss from expected and observed features in the context of surveillance. See Table 1 for more information about the variables.

Table 1: Description of variables used in computing minimum cross entropy

Variables	**Description**
Eigenfaces	Feature extraction : faces are represented as vectors in high dimensional space, and most significant features are extracted from a set of images. Using eigenfaces help to decompose a set of face images into smaller set of representative 'eigenfaces' that capture the most important variations in the face images.
Fisherface	Feature extraction : Fisher's Linear Discriminant Analysis (FLDA) is used to find combination of features and extract discriminative features that maximized the separation between different classes which minimizing variation within each class.
Facial coordinates	Through combination of several features of nose, eye, and mouth landmarks. Coordinates approaches of X,Y,Z for facial landmarks are standardized according to translation, rotation, and scaling factor.

A. Mahalanobis Distance

From a geometric perspective, the standard Euclidean distance is not utilized in the experiment since it does not take into account the correlation between highly correlated variables. In the context of face recognition, the Euclidean distance can lead to substantial result, since it does not consider correlation among pixels (Aly et al., 2008). Mahalanobis developed ways to measure distance between groups in terms of multiple characteristics is used. Mahalanobis distance played a vital role in differentiating characteristics in various different fields, including anthropology, clustering, classification, and image processing (Zhang et al., 2020). Through several works and studies, Mahalanobis gains better performance when it comes to pattern recognition problems. Therefore, an alternative approach is adopted to scale the contribution of each variables to distance value according to variability of each variables. Hence, the adoption of this metric is absolute necessary in some features, which, are:

$$D^2 = (x - m)^T C^{-1} (x - m)$$

Where:

D^2 = Mahalanobis distance
X = Vector of data
M = vector of mean values of independent variables
C^{-1}= Inverse covariance matrix of independent variables

In context to calculating minimum cross entropy between distributions, features pre-processing in Eigenfaces, Fisher faces and facial coordinates is performed before computing Mahalanobis distance. Steps required to be performed are explained in Table 2 and Table 3.

Table 2: Features pre-processing and Mahalanobis distance computation in Eigenfaces and Fisherfaces

Eigenfaces	FisherFaces
Get the standardized coordinates for the expected and observed Normalize the distribution for the standardized coordinates using PowerTransformer (Datalab, 2022) Calculate mean-centered of the expected and observed normalized and standardized coordinates Compute eigenfaces for the expected mean-centered (steps) : - Covariance matrix is computed from the expected mean-centered distribution - Perform eigen decomposition towards the covariance matrix - Select highest 10 eigen values with their corresponding eigenvectors for eigen faces Use the selected eigenvectors to project mean-centered of expected and observed into a new feature subspace to become the new expected distribution (P) and observed distribution (Q) Calculate the Mahalanobis distance 1. Distance between P and P (expected distribution) of each variable, with X = P ● Calculate covariance matrix of P^T ● Calculate mean from P ● Inverse of covariance matrix of P^T ● Calculate the Mahalanobis Distance of P and P 2. Distance between P and Q (observed distribution) of each variable, with X = Q ● Calculate covariance matrix of P^T ● Calculate mean from P ● Inverse of covariance matrix of P^T ● Calculate the Mahalanobis Distance of P and Q Calculate cross entropy between expected and observed Select minimum value, representing lowest loss between expected and observed Mahalanobis distance.	Get the standardized coordinates for the expected and observed Normalize the distribution for the standardized coordinated using PowerTransformer (Datalab, 2022) Perform necessary computations for specific requested features : - Covariance matrix is computed from the expected distribution - Perform eigen decomposition towards the covariance matrix - Select highest 10 eigenvalues with their corresponding eigenvectors for eigenfaces and Use the selected eigenvectors to project the two distributions (normalized and standardized coordinates) to become to the new expected distribution (P) and observed distribution (Q) into a new feature subspace Calculate the Mahalanobis distance 3. Distance between P and P (expected distribution) of each variable, with X = P ● Calculate covariance matrix of P^T ● Calculate mean from P ● Inverse of covariance matrix of P^T ● Calculate the Mahalanobis Distance of P and P 4. Distance between P and Q (observed distribution) of each variable, with X = Q ● Calculate covariance matrix of P^T ● Calculate mean from P ● Inverse of covariance matrix of P^T ● Calculate the Mahalanobis Distance of P and Q Calculate cross entropy between expected and observed Select minimum value, representing lowest loss between expected and observed Mahalanobis distance.

Table 3: Features pre-processing and Mahalanobis distance computation in Facial Coordinates

Facial Coordinates
Get the standardized coordinates for the expected and observed Normalize the distribution for the standardized coordinates using PowerTransformer Normalized and standardized distributions will become expected distribution (P) and observed distribution (Q) Calculate the Mahalanobis distance Distance between P and P (expected distribution) of each variable, with X = P ● Calculate covariance matrix of P^T ● Calculate mean from P ● Inverse of covariance matrix of P^T ● Calculate the Mahalanobis Distance of P and P 5. Distance between P and Q (observed distribution) of each variable, with X = Q ● Calculate covariance matrix of P^T ● Calculate mean from P ● Inverse of covariance matrix of P^T ● Calculate the Mahalanobis Distance of P and Q Calculate cross entropy between expected and observed Select minimum value, representing lowest loss between expected and observed Mahalanobis distance.

Mahalanobis distance serves as a valid approach to measure any dissimilarity between distributions in multi-dimensional space, considering correlations between variables and scales them appropriately.

IV. EXPERIMENTAL RESULTS

Our minimum cross entropy calculation is implemented and tested on several variables. Some results are encouraging. A brief discussion about the results will be explained in the next section.

A. *Minimum Cross Entropy in Eigenfaces, Fisherfaces, and Facial Coordinates*

We applied our method using the covariance matrix of the expected distribution in the *Fisherfaces and Eigenfaces.* With differing approaches of each variables, we have computed its Mahalanobis distance and their cross entropies. Figure 1 illustrates cross entropy comparisons between each of three variables, namely, Eigenfaces, Fisherfaces, and Facial Coordinates.

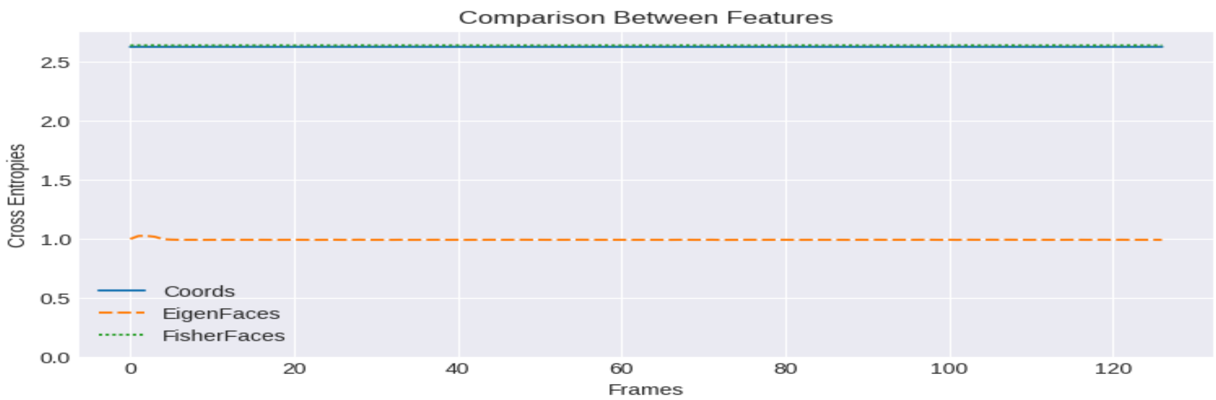

Fig. 1: Cross entropy comparison between Eigen Faces, Fisher Faces, and Facial Coordinates

Through our experiments, we have found that eigen faces resulted in the lowest value of cross entropy, meaning that this variable best accurately predicts the expected and observed distribution of images, with value approximately around 1.0, closest to the identicality of both distributions. The two other variables, Fisher faces and facial coordinates have similar values of entropy, which are around 2.6. Though unsatisfactory, a reason for this would be incompatibility of features, and feature extraction method might not be well suited for computing minimum cross entropy through Mahalanobis distance. While a high cross-entropy value may suggest that our method is not performing optimally, several reasons can be deduced from our experience. Facial coordinates have resulted in high amount of abstract data which results in higher complexity of processing it through our computational methods. Fishers faces, on the other hand, needs higher processing attention due to the need of trying to find better classification. In our case, the dimension of projection in face space is not as compact as Eigen face, which results in larger storage of face and thus more processing time in recognition. For further elaboration, experimental accuracy is computed. Accuracy is calculated based on how many true ground truth value is returned after several frames in a one minute video. In our case, the ground truth value is the expected image we have in our database, which will be given as Index 0. In a one minute video, accuracy is calculated based on how many Index 0 is returned out of several indexes present in the video. For more information, see Table 4.

Table 4: Accuracy results on each of variables

Features	Accuracy
Eigenfaces	Accuracy: 98% Min cross entropy: 0.7 - 0.92 *The model is very stable unless attempted on X and Y axis rotation.*
Fisherfaces	Accuracy: 40% Min cross entropy: around 2.6 *Took a long time to compute. The model is not stable when handling rotation on x-axis and y-axis*
Facial Coordinates	Accuracy: 16% Min cross entropy estimates: 2.62 *The model is stable. Several challenges on distinguishing between different person, but returned similar entropy values.*

ISSN No:-2456-2165

V. SENSITIVITY ANALYSIS

In our optimization problem in minimizing cross entropy, we have implemented sensitivity analysis technique to measure how responsive our output model to variations in input variables. Several facial expressions, altered conditions, as well as rotation in X,Y,Z axis have been executed to examine the stability of our model and, if any, changes to our output model. We trialed using the Eigenfaces feature, and several results from our sensitivity analysis technique exhibit a relatively stable model in our optimization problem (see Appendix). To elaborate, Figure 3 shows steady levels of cross entropies with changing facial expressions. Each experiment involved initially setting a default facial expression in early frames and progressively transitioning to the desired facial expression (see Figure 2). Through Figure 3, we may conclude that cross entropies do not exhibit significant fluctuations with changing from default expression to angry expression. This finding suggests that our model could predict stable cross entropies in response to varying facial expressions. Similar results can be seen from different facial expression, and differing conditions such as considerations of mask and glasses. However, it seems like our system is not sufficient enough to handle rotations in X and Y axis. On a good note, a relativelystable cross entropies are computed in the Z axis (see Appendix).

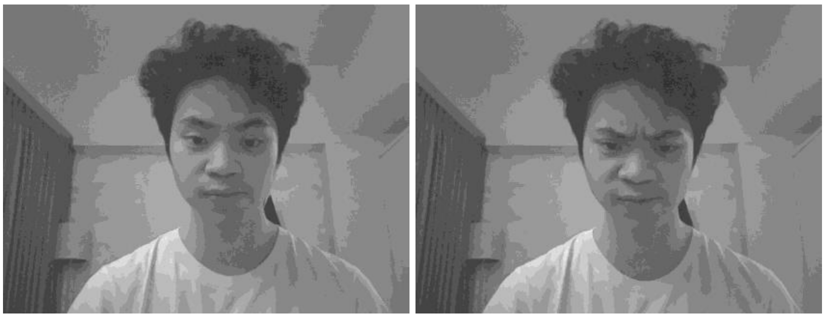

Fig. 2: Default facial expression in frame 0 progressively changes to an angry facial expression in frame 4

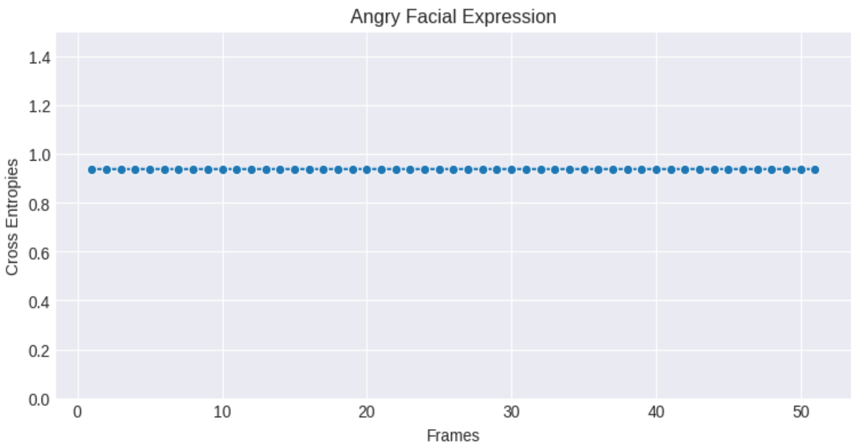

Fig. 3: Stable cross entropies from frame to frame resulted from angry facial expression

VI. CONCLUSION

Surveillance and security are absolutely essential in handling crimes and aid for suspect identification. With the rapid development of technology, manual processes of handling thousands of data needs to be transitioned into automatic method with integration of new technologies. Therefore, our model serves as a solution to many investigators having to identify suspect manually into identifying suspect in the count of minutes. Through our optimization model of minimum cross entropies, we concluded that Eigenfaces is the most reliable feature in our face recognition system in suspect identification. Our sensitivity analysis technique has provided evidence of the stability of our model in the optimization problem. The model demonstrates consistent cross entropy levels across various facial expressions, conditions, and axis rotations.

Several improvements can be made through our project. In the near future, we will consider augmenting our training dataset to include more diverse samples. Due to limited resources, our training dataset is very minimal. Having more variations in our training set of images can help to improve the performance of our model. In addition, we will explore alternative feature engineering techniques that may better capture underlying patterns in our dataset. Several feature extraction algorithms such as Local Binary Pattern Histogram (LBPH), Scale-Invariant Feature Transform (SIFT), and Convolutional Neural Networks (CNN) can be trialed to see the suitability of these features in our optimization problem of the face recognition system.

In terms of accuracy handling of our facial recognition software system, in the near future we will attempt to augment our training images dataset with a diverse set of face images with various pose variations. With larger datasets of different rotations, we believe that our facial recognition software system will be robust to handle axis rotations.

REFERENCES

[1.] Muftić, E. B. & N. (2021, August 27). Facial recognition: An introduction. Institute for Internet and the Just Society. https://www.internetjustsociety.org/cosmonaut/facial-recognition-an-introduction

[2.] Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. Journal of Cognitive Neuroscience, 3(1), 71–86. https://doi.org/10.1162/jocn.1991.3.1.71

[3.] Anggo, M., & La Arapu. (2018). Face recognition using Fisherface method. Journal of Physics: Conference Series, 1028, 012119. https://doi.org/10.1088/1742-6596/1028/1/012119

[4.] Brownlee, J. (2020, December 22). A gentle introduction to cross-entropy for Machine Learning. MachineLearningMastery.com. https://machinelearningmastery.com/cross-entropy-for-machine-learning/#:~:text=Cross%2Dentropy%20is%20a%20measure%20of%20the%20difference%20between%20two,encode%20and%20transmit%20an%20event.

[5.] Gentile, G. (2023, March 27). How Technology is Revolutionizing the Field of Security Services. SecurityRI. https://www.securityri.com/revolutionizing-the-field-of-security-services/

[6.] Insider, I. (2021, January 1). Role of CCTV Cameras : Public, Privacy and Protection. IFSEC Insider | Security and Fire News and Resources. https://www.ifsecglobal.com/video-surveillance/role-cctv-cameras-public-privacy-protection/

[7.] Hildreth & Rueda. (n.d.). What factors can make eyewitness testimony unreliable?: Hildreth & Rueda: Austin, Texas. What factors can make eyewitness testimony unreliable? https://www.youraustinattorney.com/articles-blog-posts/what-factors-can-make-eyewitness-testimony-unreliable/

[8.] Aly, S., Tsuruta, N., & Taniguchi, R.-I. (2008). Face recognition under varying illumination using mahalanobis self-organizing map. Artificial Life and Robotics, 13(1), 298–301. https://doi.org/10.1007/s10015-008-0555-z

[9.] Y. Zhang, B. Du, L. Zhang and S. Wang. (2020). A low-rank and sparse matrix decomposition-based Mahalanobis distance method for hyperspectral anomaly detection. IEEE Transactions on Geoscience and Remote Sensing 220 (2016), 1376– 1389.

[10.] Datalab, A. (2022, March 28). Understanding power transformer. Medium. https://medium.com/@analyttica/understanding-power-transformer-2bf6e18717b3

APPENDIX

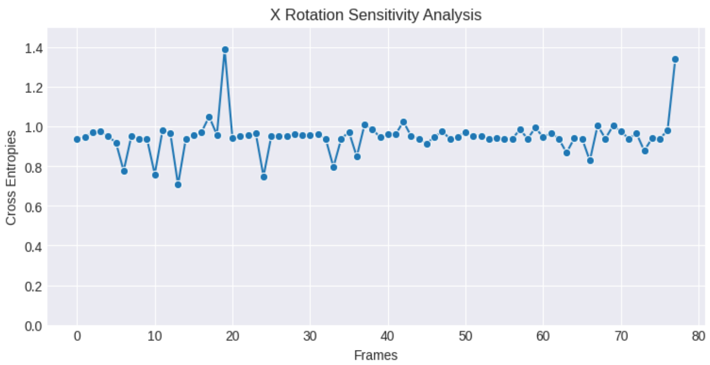

Fig. 4: Cross entropies from frame to frame X-Rotation

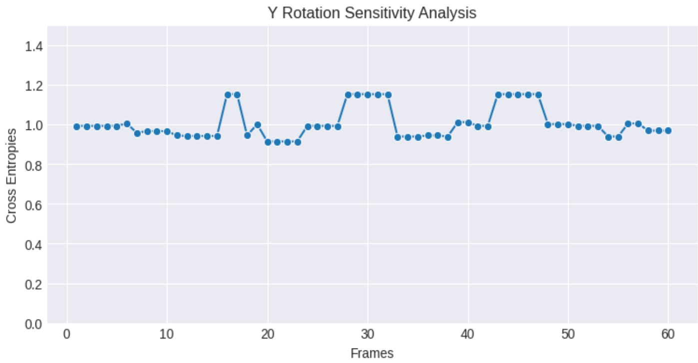

Fig. 5: Cross entropies from frame to frame Y-Rotation

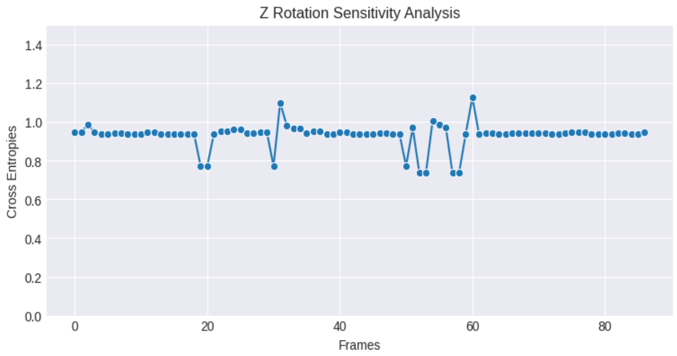

Fig. 6: Cross entropies from frame to frame Z-Rotation

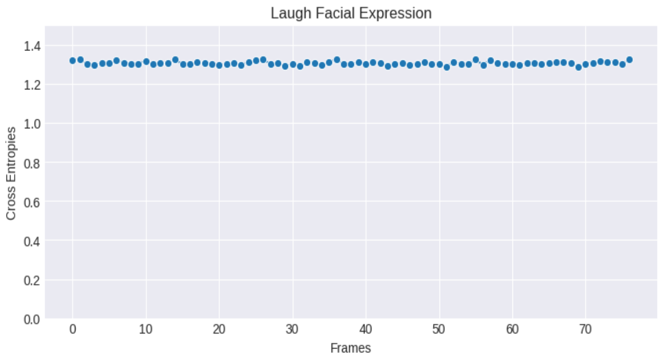

Fig. 7: Cross entropies frame to frame laugh expression

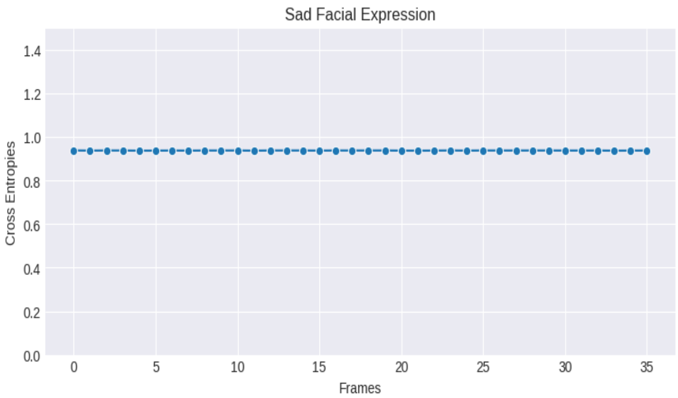

Fig. 8: Cross entropies frame to frame sad expression

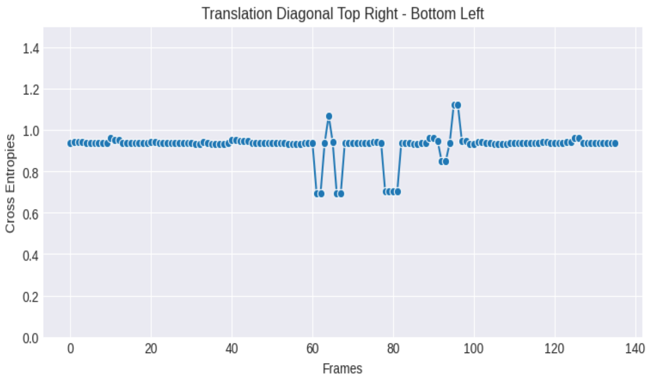

Fig. 9: Cross entropies frame to frame diagonal translation

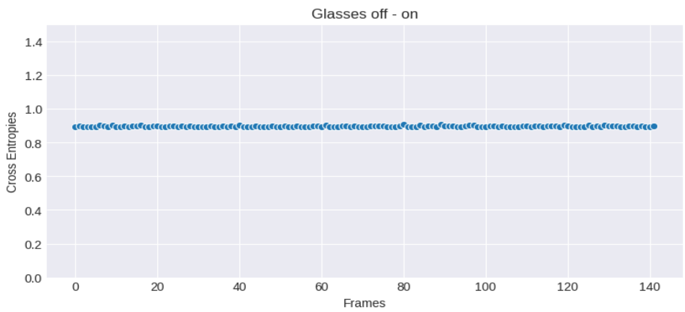

Fig. 10: Cross entropies frame to frame glasses off and on

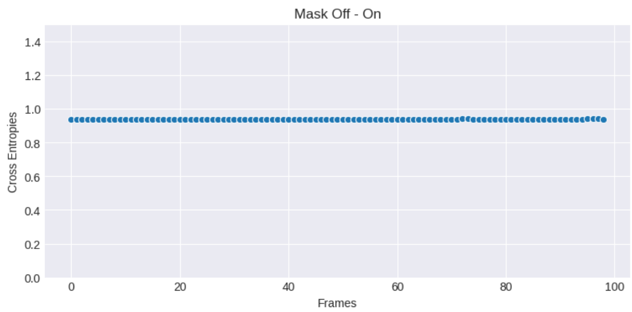

Fig. 11: Cross entropies frame to frame mask off and on