Hybrid techniques based on deep learning and mathematical approach for lips movements detection

Nagham salim[1]

[1]Department of computer Science, collage of science, mosul University, Ninva, Iraq
Email: naghamallella@gmail.com

Khalil al-saif[2]

[2]Department of Computer Science, collage of mathematic and computer science ,mosul University, , Ninva, Iraq

ABSTRACT

Lips movement make new approach for human relation ships specially for hearing impaired persons, so a lot of researches are done to make service for them via a lot of IOT application. In this research deep learning was adopted for tracking lips movements for speech understanding. In addition to the deep learning a mathematical approach are applied via curve fitting using polynomial equation with Euclid's distance which applied in different words and different genders. The proposed algorithm show the hybrid approach has more benefit compared with traditional technique. The data set used its locally generated with special Arabic selected word.

Keywords: lips tracking, curve fitting, landmarks , Dlib based CNN , Face detection

1. Introduction

The analysis of Visual speech (VSS) research began since the beginning of the twentieth century, and because of its applications in several fields such as virtual reality, the field of language communication between humans, the conceptual field and other applications, which led to the rapid development of this technology, and the important element in visual speech analysis is the movement of the lips [1].

The process of extracting the feature of the lips or tracing the lips is a complex process with the same problems of face detection, such as the difference between people, the variation of lighting, etc. Feature extraction of lips seems to be very sensitive to negative circumstance parameters. For example, moustache is easy to believe that it is the upper lip and teeth when the absence of a sharp contrast between the lips and the face which complicate the extraction of the lip feature. Using knowledge of lip color or shape in order to identify and track lips is considered as one of the modern approaches, when color differentiation is an effective technique to locate the lips. [2].

The analysis of Visual speech (VSS) techniques depend on the lips movement, as the extracted information is two-dimensional and contains visual information that is more repetitive. therefore, dealing with visual speech is more difficult than dealing with the one-dimensional sound [3].

In this research, we relied on the analysis of the lips movement, and the movement path was represented by the adoption of a polynomial function, in addition to the adoption of measurements for the purpose of calculating the distances for internal and external mouth opening, and for the purpose of discovering the face area, the Dlib CNN technique was adopted, and use Detector in Dlib library for the purpose of

discovering the key points (landmarks) of the face area.

2. Previous Work

Soukupova, Tereza and Cech,jan,2016, presented an effective algorithm for the purpose of detecting eye blinking by relying (SVM) , in addition to deriving standard quantities (EAR) the ratio of eye aspect, to determine the level of the eye aperture [4].El-Melegy, Moumen T., et. al. in their research at 2019 divided the work into several stages with the first stage they insert the image into three types of face detectors. The output of this stage which is the faces detected in the image because the features of the haar, so the detectors that rely on it to detect faces are used. Within the second stage the faces discovered in a particular detector are an entry to another detector to show its ability to detect those faces. Finally the third stage detecting the skin color area for each bounding box of the detected faces in the previous stage and the degree of the detector is measured [5].

Parra-Dominguez, Gemma S., et. al.,2021, studied the patient who is paralyzed in the face area and try to use the computer programs and systems for the purpose of detecting paralysis in the facial muscles which is important for the purposes of medical evaluation and monitoring. The proposed algorithm extracting the basic features and making special measurements on those features . It uses effective mathematical operations for the purpose of measuring the key points extracted for the face area [6].

The discovery of the face region (the interested region) were studied by Képešiová, Zuzana and Kozák, Štefan,2018,, which is a specific part of the whole organism, requires finding the size and location of all parts (eye area, nose, mouth, eyebrows) that make up the face area. This algorithm deals with faces that are forward-facing and online. they mainly extracts the eye and mouth areas as the key features of the face area. The researcher provided an accurate description of the functions of effective face detection algorithms and accurate analysis of solutions for the purpose of reaching the best results, used (viola jones) for the purpose of discovering the face area, (casscade) for the purpose of classification, and either training the algorithm using (Adaboost) and thus this algorithm was characterized by speed in image processing with a high rate Face detection speed.

discovery of the face region (ROI), which is a specific part of the whole organism, requires finding the size and location of all parts (eye area, nose, mouth, eyebrows) that make up the face area. This algorithm deals with faces that are forward-facing and online, as it detects the face entered into the algorithm based on the patterns it has previously trained on, and mainly extracts the eye and mouth areas as the key features of the face area. As for the problem of the multiplicity of objects and overlapping each other, the solution was to store the size and location of these spaces for the purpose of comparing them and placing a certain weight for each feature in their locations [7].

Khansaa and Stanciu (2020), The researcher suggested a system for facial recognition, as he presented a system that could be adopted in a smart house or an official building It needs security intervention in general. Facial recognition is done through a series of images or video tracking, this technology distinguishes people by relying on a specific algorithm, the algorithm applied by the researcher to discover objects is Viola–Jones, the proposed system provides two-step facial recognition: the first is to detect the face from the video directly Using the computer's webcam. The second is to decide whether this face has permission to enter the building or not, by searching into an existing database[8].

Noda et al. (2014), It was used CNN as a tool to extract the properties of the lip images, and the experimental results showed that the visual properties obtained using Convolutional Neural Network (CNN) are much better than the traditional methods including principal component analysis [9].

Ahmad and Sabah, (2010), This research deals with the lip positioning of the visual speech recognition system, as it presented an efficient method for determining the position of the lips of a person in the images taken from the video. These methods based on using

YCbCr approach as a starting step to extract at least a lip's portion, then all available information will be used on lip cropped pixels, by calculate the mean for each value, and then for each pixel in the ROI, then distance of Euclid is calculated from the vector of ROI, The pixels with the least distance are considered as lip pixel, and so on for the rest of the pixels in the ROI area [10].

3. Face Detection

Finding face in images called face detection, its first and essential step in many applications such as face identification, applying facial features, and tracking face movement. The effectiveness of these tasks depends on the accuracy of face detection, as the accuracy and efficiency of face detection is one of the challenges confrontation face detection methods and the reason for this is due to several factors, including facial expressions, lighting, and image blurring in addition to many factors.

Detection the face, characterized by smaller differences in the specific part (face) compared to the detection of the whole object as the differences are larger in relation to size (the difference in several pixels is much less than the difference in thousands of pixels)[11].

In the computer vision finding face it's a problem need to apply many complex computational steps that must be implemented using a computer and adopts algorithms for the detected the face in the images[12].

Face detection one of the face applications in field of computer vision, which is considered one of the most common problems. Techniques that used in this field have evolved from traditional techniques in field of computer vision to more complex machine learning programs. The main step used in these techniques is to discover the area in image which there is a face or nothing[12]. An Arabic Visual Dataset for Visual Speech Recognition

Images sequence (frames) extracted from the video, is taken as input for the purpose of locating the face area (ROI) in these images (frames) and in order to reach this specific area in the object the face area is separated from the background area (non-face), within the detected face area (ROI) determines the locations of important features (Mouth, eyes, nose and eyebrows) This process is called extracting facial features, which are later adopted for the purpose of tracking the movement of the lips[13]. Face detection techniques divide the image into a section that contains faces (ROI) and the other contains the background, and because of the difference in terms of age and skin color, in addition to facial expressions, which are common factors between faces, as well as the different lighting conditions and characteristics of the image in terms of engineering and image quality related to the means of taking the image Or video, these factors and other challenges that will be addressed later, made the treatment of face detection a difficult process. Face detection techniques must be distinguished by their ability to detect any face under any lighting condition and any background[14].

3.1 face detection challenges:

- Unnatural facial expressions in the image (expression varation).
- Presences of Glasses, hand, facial hair, hats or any other object that may lead to the masking of the face in the image and the inevitable result of this is a reduction in the rate of face detection techniques.
- The contrast in the intensity of lighting in a single image, one part of the image may contain a high light intensity, while the intensity of the lighting is low in another part.
- Another factor that reduces the rate and accuracy of face detection techniques is the presence of different objects in the background of the image, in other words, a complex background[15].
- There is a large number of human faces in the image.
- Variation in skin color from person to person, for example, skin color in Africa is very different from skin color in America.
- In addition to the above factors, the large distance between the face and the camera, in addition to the angle of the direction of the face, is considered one

of the challenges facing the face detection processing [16].

Several methods have been proposed, such as deep learning CNN, deep learning DNN and Cascade Classifier for the purpose of solving the above problems.

3.2 Dlib based CNN:

Convolutional Neural Network (CNN) is one of the deep learning algorithms that used to analyze the images entered into the network, while the Dlib is an open software library that contains thousands of algorithms in various fields of machine learning that are used to solve complex problems.

Dlib CNN, its combination of the convolutional neural network and the dlib library, this technology uses dlib tools as well as cnn features for the purpose of face detection[17]. Dlib face detection based on CNN, in addition to the features of a convolutional neural network, this method is based on MMOD(Maximum-Margin Object Detector), the filter used to extract properties from the image is manually determined in other techniques, while in this technique only the number of filters used is determined.

3.3 How Dlib CNN dose work:

First step: face detection model must be downloaded first. filters that used for the purpose of extracting properties from the image, they are determined automatically by the model. This model is loaded once for executing the algorithm, model exciting in the model weights file.

second step: The weights file that was loaded in the previous step will be entered at this stage (dlib cnn initialization stage for the purpose of detecting faces)

Third step: After configuring the detector, it is applied to the images entered into the algorithm and to identify and discover the faces in those images.

Dlib CNN is characterized by its ability to detect faces in different directions (non-frontal faces) in addition to frontal faces and blockages of faces. This algorithm is implemented in very simple steps, which is why it is easy to train, in addition to having high computational capabilities.

One of the disadvantages of this method is that it does not work with the video, so the video must be converted into frames. for the small faces in the images cannot be detected using Dlib CNN, as the smallest size of the faces that can be extracted is 80 * 80 [17].

4. The Landmarking Of Face

Discovering and identifying important points in the face are key points used for several purposes in facial treatment, such as tracking facial curves, distinguishing facial expressions, distinguishing and identifying the face and other applications. Landmarking the face is an intermediate stage that plays an important role in the advanced stages of facial treatments, as the two ends (corners) of the eye, the bottom of the nose, the edges (corners) of the nose, the area where the lips meet (the beginning and end of the lips), the beginning and end of the eyebrow are considered points face key[18]. The accuracy and robustness of determining the key points of the facial features may decrease due to some factors:

1. Determining the key points of the facial features in the correct way differ depending on the contrast of faces between individuals, and these are internal influential factors, while the external factors are obstruction, lighting contrast, camera standards, face direction and facial expressions.
2. The number of key points on the face and the required accuracy. The nose and eye area needs to be defined and with high accuracy, as it is relied upon in determining the secondary features in the face[19].

The number of key points required and percussion varies depending on the applications, as simple-level applications may need only the basic key points represented in the meeting area of the lips, the corners of the eyes and the area under the nose, in addition to the square definition that defines the face[20].

As for the applications with higher levels, they need more number of key points between 20-30 and 60-80, applications that need more accuracy determine the area of the nose and

eyes more accurately (where they are the most prominent in the face and the most reliable), which are used later to find the points Secondary key in the face less prominent. The accuracy of the key points in the face is directly proportional to their number, as increasing the number of points to 68 key points in the face gave an increase in accuracy by 50%[21]. For the purpose of estimating the features of the face, an algorithm (estimation the face orientation) is used. Finding 68 key points in the face is the basic idea of this algorithm[22].

When applying machine learning algorithms, it starts to automatically distinguish the key points of the face and find the interested region (ROI) and the algorithm adopted in this research (detector for facial key point in Dlib library), where it works to find 68 key points of the face and represents the coordinates of these points in pixels, It is represented by the figure(1),[23].

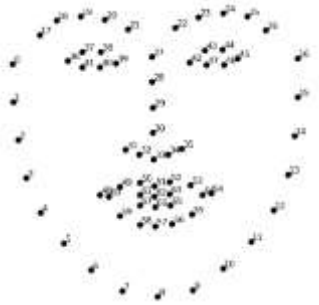


Figure 1. all 68 face key point (landmarks)

5. Curve Fitting

It is used in data analysis and helps in predicting the analysis and clarifying graphically the related data points, and distinguish whether a linear or non-linear model. Adaptation of curves may be used for the purpose of smoothing data and improving the curve and is concerned with finding centers along the curve. It checks the correlation between the dependent variables and the independent variables in order to find a good suitable model for that data. Numerical structure in fitting curves has a higher amount than relationship, using a type of data entry that formulates a mathematical function. Uses curves alignment to match a suitable model or curve to present data points is a primary necessity for many areas such as computer graphics, image processing and data mining. the type of models to fit the curve are linear, polynomial of various degrees, power fit, logarithmic curve fit and non-linear curve fits, curve fitting not only works on disorganized data in different models, but also performs different tasks for the purpose of reducing noise, creating mathematical relationship between variables and evaluating attributes between data models[24].

5.1 Polynomial regression:

For nonlinear models, polynomials can be used for their convenience, although many models in actual problems are not suitable for linear models, the effect of fit is weak. Although polynomials have good error training for trained data, it is easy to generate over-fitting. Mainly the curve will fit much better the training data. Particularly when the model order increases, However, for test data instead training data, higher-order models often do not have a good favourable effect on low-order models[24].

Polynomial regression typically has two solutions: the normal equation. polynomial regression model and the gradient descent method distributed over linear model, quadratic curve model, cubic curve model, etc. based on the number of highest items. The linear model as in equation (1)[25]:

$$Y=a + bx+ cx2 +dx3 + ex4\ldots\ldots + etc. \ldots\ldots\ldots (1)$$

The goal of polynomial regression is to determine the values of the parameters that make the curve fit the data points in the best way. The order of the equation (number of coefficients) must be determined. When the number of coefficients is two, then the equation

describes a line and the polynomial regression is identical to the linear regression, while when the number of coefficients is three or above the equation then Describes a curve, the more coefficients the more elastic curve it creates[26].

6. **Method:**

- **Face region detection:**

The Dlib library contains many functions used to force face region detection, and one of these functions is Dlib MMOD CNN (dlib.cnn-face-detection-modle-v1) function, This function is distinguished from others its more accurate when it works to discover the face area in all directions, but it needs a longer period of time due to complex calculations. show Execution of Dlib CNN on the entered Frames is a set of images for detected face area with the extension (.jpg), as show in figure(2).

Figure 2.a. Original frame

Figure 2.b. face detection took 10.6450 seconds using Dlib CNN

- **Determining the land marks of the face area:**

At this stage and for the purpose of determining the main key points of the face, use Dlib which is one of the machine learning tools and was designed for the purpose of extracting the key points of the face because it has tools to extract and train the (68) key point for the face area. for discovery land marks of the face region is a part methods of the shape predictor (the shape of predictor), which is located inside the Dlib library, its work to determine the locations of the main key points of the face features and to estimate the coordinates (x,y) of these points and for the purpose of drawing the final map structure of the face, and each face image input to this algorithm, the key point finder (facial landmarks detection) is applied. figure(3), show localize pert from face to extracted mouth image.

Figure 3. Localize part from face to extracted mouth image

- **Measuring The Mouth Opening:**

1. Calculate The Distance To The Key Points Of The Inner Mouth Opening Which Are Represented By The Dots From (60 - 67), As In Formula (2)[27].
2. Calculate The Distance To The Key Points Of The Outer Mouth Opening Which Are Represented By The Dots From (48 - 59), As In Formula (3)[27].
3. Calculation Of Euclid's Horizontal And Vertical Distance For The Outer And Inner Opening Of The Mouth, By Take The Average Distance For These Key Points, Then Calculate The Aspect Ratio, As In Formula (4).

Mouth_Vertical= Mean(Dis(P61,P67), Dis(P62,P66), Dis(P63,P65)) *……..(2)*

Mouth_Horizontal= Dis(P60,P64) *……..(3)*

Mouth_Inner Ratio= Mouth_Vertical/ Mouth_Horizontal .....*(4)*

In The Same Steps, The Ratio Of The Height And Width Of The External Mouth Opening Is Calculated, But Based On The Key Points Associated With The External Mouth Opening. Figure(4), Show Horizontal And Vertical For Outer, Inner Mouth Distance.

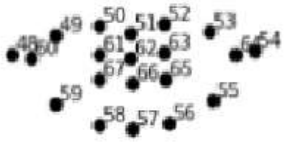
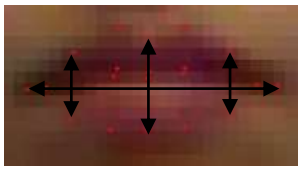
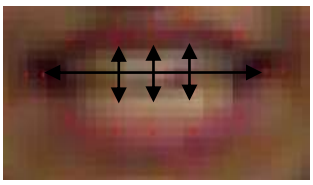


Figure 4.a. Landmark points for mouth region



b. Horizontal and vertical outer mouth distance



c. Horizontal and vertical inner mouth distance

Table 1. Inner and outer mouth measurment		
Frame	Outer ratio mesurment	Inner ratio measurment
Frame1	0.4497	0.07157
Frame2	0.4531	0.0716
Frame3	0.4412	0.0672
Frame4	0.4379	0.0666
Frame5	0.4513	0.0625
Frame6	0.4790	0.1624
Frame7.....frame11	.....	.....
Frame12	0.4111	0.0713
Frame13	0.4054	0.0674
Frame14	0.3923	0.0625

Table 1. Show the result for inner and outer close and open mouth during speaker talk one words (مرحبا), this measurement applied to 14 frames represent lips movement.

- **Lips preprocessing:**

After extracting the interested region (mouth area) and for the purpose of extracting the edge of the lip region, the lips region is processed using the chromatic eraser (HSV, hue and saturation of value) and then Thresholding technique is applied to the extracted color image, the output of this stage is passed on the Morphological process and finally for the purpose of extracting the edge of the lips Canny Filter is applied, Because of its ability to detect a larger number of symmetric edges. After passing a window whose size depends on the size of the image entered into processing, the extracted edge of the lower lip region is taken for the purpose of extracting the points of that edge, and we fit these points to the extracted lower lip edge with a polynomial curve. Thus, we have obtained coefficients for the polynomial equation of degree (4) and for the purpose of obtaining the best fit we use the function poly1d and the next step is to find the Mean for these coefficients.

the polynomial function is expressed in the format, wher a, b, c, d, e, f, are the polynomial coefficient value, Table(2) represent polynomial coefficient for different degree that applied on mouth edge, where figure (5) show plotting polynomial coefficient.

$$Y=a + bx + cx2 + dx3 + ex4 + fx5$$

a =3.053 , b =-5.562, c =1.915, d =3.064, e =5.88

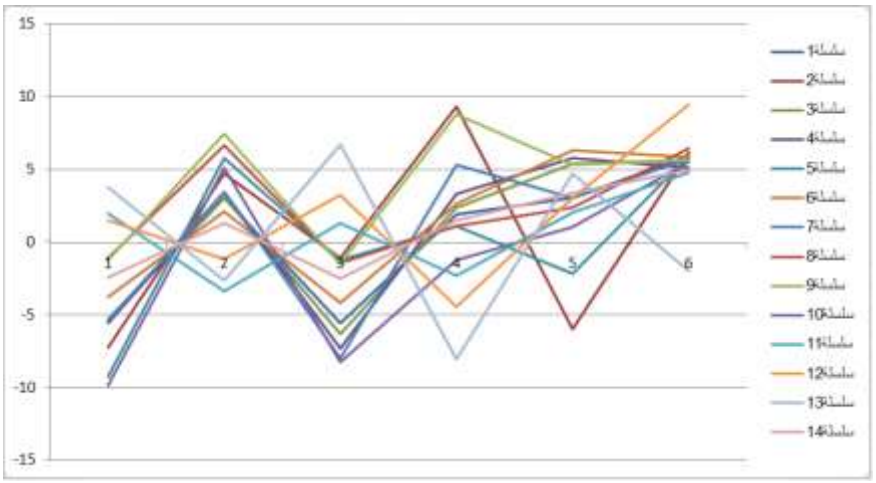
Mean=6.408929, represent mean for polynomial coefficient extracted from (14) frame to the lower lips movement said (مرحبا)word from first video

Mean=5.980357, represent mean for polynomial coefficient extracted from (14) frame to the lower lips movement said (مرحبا)word from second video

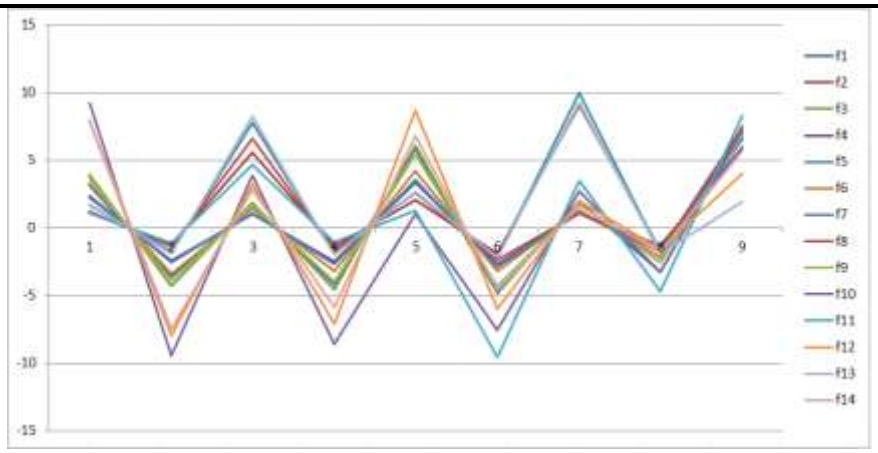
Table 2. Polynomial coefficient with different degree apply on lower mouth edge

	F1	F2	F3	F4	F5……F10	F11	F12	F13	F14
Polynomial Degree (5)	-5.391	-7.226	-5.302	-5.579		1.996	1.49	3.795	-2.439
	3.053	4.719	3.122	3.414		-3.384	-1.152	-2.634	1.332
	-5.562	-1.098	-6.336	-7.298		1.33	3.259	6.74	-2.482
	1.915	9.31	2.368	3.282		-2.33	-4.446	-8.083	1.506
	3.064	-5.974	5.39	5.765		2.018	3.073	4.7	3.348
	5.898	6.177	5.659	5.14		4.785	9.475	-1.971	5
Polynomial Degree (8)	3.164	1.176	3.949	2.245		1.092	7.916	1.778	7.963
	-3.576	-1.238	-4.273	-2.382		-1.109	-7.899	-1.882	-7.463
	1.671	5.577	1.905	1.05		4.651	3.243	8.279	2.865
	-4.161	-1.396	-4.514	-2.484		-1.038	-7.054	-1.955	-5.832
	5.922	2.097	6.095	3.395		1.323	8.69	2.673	6.802
	-4.775	-1.874	-4.659	-2.67		-9.537	-5.961	-2.122	-4.561
	1.996	9.061	1.84	1.092		3.536	2.002	9.256	1.609
	-3.266	-1.691	-2.552	-1.307		-4.649	-1.527	-1.894	-1.779
	7.545	7.001	6.998	6.003		8.283	4.056	1.956	5.697
Polynomial Degree (15)	-6.743	-1.559	5.129	-4.802		2.046	4.893	1.138	-4.43
	1.331	2.572	-9.851	8.861		-3.949	-9.55	-2.228	5.598
	-1.177	-1.833	8.415	-7.346		3.441	8.479	1.985	-2.511
	6.159	7.255	-4.23	3.611		-1.789	-4.521	1.985	1.852
	-2.116	-1.663	1.397	-1.17		6.1878	1.616	3.845	2.839
	5.013	1.833	-3.218	2.627		-1.498	-4.093	-9.856	-1.43
	-8.345	9.074	5.374	-4.181		2.612	7.5598	1.848	3.602

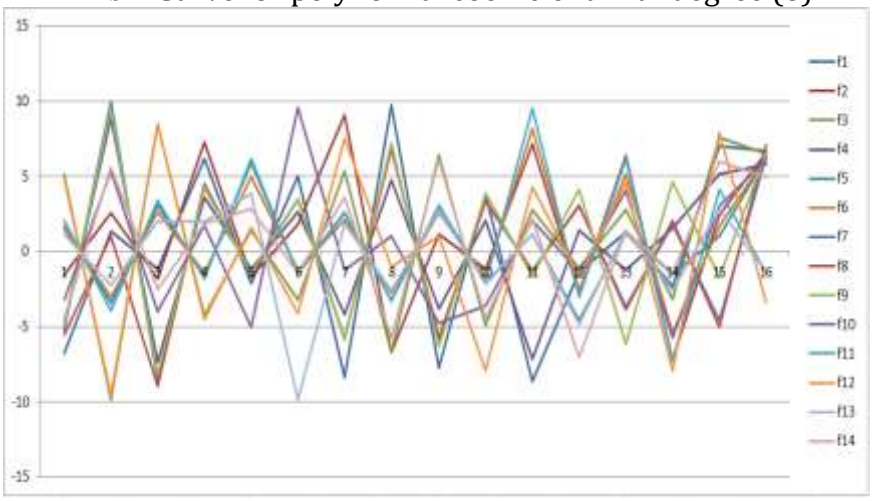
	9.73	-6.713	-6.718	4.734		-3.314	-1.033	-2.576	-5.729
	-7.708	1.163	6.469	-3.77		3.059	1.0497	2.68	6.12
	3.811	-1.149	-4.863	2.05		-2.034	-7.894	-2.076	-4.454
	-8.652	7.137	2.787	-7.177		9.53	4.337	1.183	2.186
	-1.358	-2.777	-1.126	1.438		-3.034	-1.701	-4.86	-7.008
	1.283	6.344	2.78	-1.246		6.18	4.582	1.387	1.371
	-2.31	-7.182	-3.131	1.546		-7.252	-7.95	-2.595	-1.42
	6.99	2.439	7.567	5.122		4.14	7.914	2.84	5.983
	6.675	6.597	6.477	5.767		-1.397	-3.347	-1.366	4.968



a. Curve for polynomial coefficient with degree (5)



b. Curve for polynomial coefficient with degree (8)



c. Curve for polynomial coefficient with degree (15)

Figure 5. Represent curve for lower lips edge polynomial coefficient

7. Conclusion

In this research, a polynomial function of degree four was used in order to derive the movement of the lower part of the lips to imitate the lips during the pronunciation of the word, After deriving the equation for a number of frames (14 frame) the mean was found for these polynomial equation coefficients, and when comparing the arithmetic mean of two videos taken at the same time period and for the same word (مرحبا) the result was very close.

For the purpose of discovering the face area, Dlib CNN was adopted, which gave accuracy in the results and ease in implementing the process of discovering the face in various angle, but the discovery process needs a period of time. as for the purpose of finding main key points for the mouth region (the interstate region), which is part of the face region, use detector for facial key point in Dlib library, which gave accurate results in finding the key points of the mouth area in each picture during the opening and closing of the mouth from the beginning of the pronunciation of the word to the end of that word.

The number of frames extracted from the video (14 frames), for a speaker speaking one word, and Euclid's distance for mouth movement was calculated to measure the internal and external mouth dimensions from the aspect ratio of the mouth opening for each frame

Reference:

1. Salleh, Siti Salwa. et al. 2009. 'Synthesis Visual Speech using Modified Polynomial Motion Curve', IJCSNS International Journal of Computer Science and Network Security, VOL.9 No.2, February 2009.
2. Duy Nguyen. et al. 2006. 'Real-Time Face

Detection and Lip Feature Extraction Using Field-Programmable Gate Arrays', IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, VOL. 36, NO. 4, AUGUST 2006

3. M, Hao. et al. 2020. 'A Survey of Research on Lipreading Technology', November 9, 2020, date of current version November 20, 2020. Digital Object Identifier 10.1109/ACCESS.2020.3036865, IEEE ACCESS.
4. T, Soukupova and J, ´ Cech. 2016. 'Real-Time Eye Blink Detection using Facial Landmarks' , 21st Computer Vision Winter Workshop Luka Cehovin, Rok Mandeljc, Vitomir ˇ Struc (eds.) ˇ Rimske Toplice, Slovenia, February 3–5, 2016.
5. M, T. El-Melegy. et al. 2019. 'Optimal Score Fusion via a Shallow Neural Network to Improve the Performance of Classical Open Source Face Detectors', Optimal Score Fusion via a Shallow Neural Network to Improve the Performance of Classical Open Source Face Detectors. DOI: 10.5220/0007691206600667 In Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2019), pages 660-667 ISBN: 978-989-758-354-4.
6. G, S. Parra-Dominguez. Et al. 2021. 'Facial Paralysis Detection on Images Using Key Point Analysis', Appl. Sci. 2021, 11, 2435. https://doi.org/10.3390/app11052435.
7. Z, Képešiová and Š, Kozák. 2018. 'An Effective Face Detection Algorithm', Conference Paper · January 2018 DOI: 10.1109/CYBERI.2018.8337566,https://www.researchgate.net/publication/324554906.
8. K. Dheyaa Ismael1 and S. Irina. 2020." Face recognition using Viola-Jones depending on Python", Indonesian Journal of Electrical Engineering and Computer Science Vol. 20, No. 3, December 2020, pp. 1513~1521.
9. K. Noda. Et al. 2014, "Lipreading using convolutional neural network," in Proc. Conf. Int. speech Commun. Assoc., pp. 1149–1153.
10. Ahmad B. A. Hassanat, Sabah Jassim. 2010. "Color-based Lip Localization Method", CONFERENCE PAPER in PROCEEDINGS OF SPIE - THE INTERNATIONAL SOCIETY FOR OPTICAL ENGINEERING · APRIL 2010 Impact Factor: 0.2
11. S, Minaee. et al. 2021, 'Going Deeper Into Face Detection: A Survey',arXiv:2103.14983v1 [cs.CV] 27 Mar.
12. Md. Hasan. Et al. 2021. 'Human Face Detection Techniques: A Comprehensive Review and Future Research Directions', Electronics 2021, 10, 2354. https://doi.org/10.3390/electronics10192354.
13. L. A. Elrefae. Et al. 2019, 'An Arabic Visual Dataset for Visual Speech Recognition', Procedia Computer Science 163, 400–409.
14. Omaima N. A. AL-Allaf. 2014. 'REVIEW OF FACE DETECTION SYSTEMS BASED ARTIFICIAL NEURAL NETWORKS ALGORITHMS', The International Journal of Multimedia & Its Applications (IJMA) Vol.6, No.1, February.
15. Luo D, Wen G, Li D, Hu Y, Huan E. 2018. 'Deep-learning-based face detection using iterative bounding-box regression', Multimed Tools Appl, https://doi.org/10.1007/s11042-018-5658-5.
16. Kumar. Et al. 2019. 'Face detection techniques: a review', Article in Artificial Intelligence Review · August 2019 DOI: 10.1007/s10462-018-9650-2.
17. Jadhav. Et al. 2021. 'Survey on Face Detection Algorithms', International Journal of Innovative Science and Research Technology, Volume 6, Issue 2, February , ISSN No:-2456-2165
18. Johnston and P. de Chazal. 2018. 'A review of image-based automatic facial

landmark identification techniques', Johnston and Chazal EURASIP Journal on Image and Video.
19. H. OUANAN. Et al. 2016. 'Facial Landmark Localization: past, present and future', 978-1-5090-0751-6/16/$31.00IEEE, https://www.researchgate.net/publication/312111626.
20. S. Badura and M. Mokrys. 2015. 'Feature extraction for automatic lips reading system for isolated vowels', Human computer interaction and interface, visualization and natural language processing, eISSN: 1339-9144, cdISSN: 1339-231X, ISBN: 978-80-554-1002-9
21. O. Çeliktutan. Et al. 2013. 'A comparative study of face landmarking techniques', Çeliktutan et al. EURASIP Journal on Image and Video Processing 2013, 2013:13.
22. N. Boyko. Et al. 2018. 'Performance Evaluation and Comparison of Software for Face Recognition, based on Dlib and Opencv Library', IEEE Second International Conference on Data Stream Mining & Processing August 21-25, 2018, Lviv, Ukraine.
23. S. Monini. Et al. 2021. 'An objective, markerless videosystem for staging facial palsy ', European Archives of Oto-Rhino-Laryngology (2021) 278:3541–3550 https://doi.org/10.1007/s00405-021-06682-z, Content courtesy of Springer Nature, terms of use apply. Rights reserved.
24. V. P, D. and R. Rao. 2016. 'Machine Learning Techniques on Multidimensional Curve Fitting Data Based on R- Square and Chi-Square Methods', International Journal of Electrical and Computer Engineering (IJECE) Vol. 6, No. 3, June, pp. 974 ~ 979.
25. X. Wan. 2019. 'The effect of regularization coefficient on polynomial regressionk', IOP Conf. Series: Journal of Physics: Conf. Series 1213 (2019) 042054.
26. H. J. MOTULSKY and L. A. RANSNAS. 1987. 'Fitting curves to data using nonlinear regression: a practical and nonmathematical review', FASEB Journal.
27. L. Chen. Et al. 2021. 'Driver Fatigue Detection Based on Facial Key Points and LSTM', Security and Communication Networks, Volume 2021, Article ID 5383573, 9 pages.