International Journal of Engineering Technologies and Management Research

IJETMR

A Knowledge Repository

FACE RECOGNITION WITH HYBRID TECHNIQUES

Prashant Kushwah [*1]
[*1] ASET, Amity University Madhya Pradesh, India

Abstract:
Face recognition framework is still in test by numerous applications particularly in close perception and in security frameworks. Generally all utilizations of face recognition utilize enormous information sets, making challenges in present time preparing and effectiveness. This paper contains a structure to enhance face recognition framework which have a few phases. For good result in face recognition framework a few upgrades are critical at each stage. A novel plan is displayed in this paper which gives the better execution for face recognition framework. This plan incorporates expanding in datasets, particularly huge datasets which are required for profound learning. Changing the picture differentiate proportion and pivoting the picture at a few edges which can enhance the recognition precision. At that point, trimming the proper territory of face for highlight extraction and getting the best element vector for face recognition finally. The last after effect of this plan will demonstrate that the given structure is able for distinguishing and perceiving faces with various postures, foundations, and appearance in genuine or present time.

Keywords: *Framework; Face Recognition; Deep Learning; Preprocessing.*

1. Introduction

Face recognition is a one of the best biometrics acknowledgment system which has been utilized generally as a part of our business field and our day by day life. It has been moreover a standout amongst the most famous research heading in the field of computer vision and a characterization issue in machine learning. Face recognition has numerous focal points like none contact, all the more well-disposed, continuous and more satisfactory look at to different biometrics Acknowledgment like Fingerprint, Iris acknowledgment, more friendly also, step Acknowledgment. The exploration of face acknowledgment has been created for around 50 years, though it is likewise have high research esteem. The advancement street of face Acknowledgment is viewed as two stages. The principal stage is utilizing the traditional technique like Eigen Face, Fisher Face and Local binary histogram (LBPH) for face acknowledgment when the 2nd stage is utilizing the profound learning technique that exceptionally celebrated lately. The traditional strategies are utilized for little dataset and can as it were conquer a few issues. Taking Eigen Face for instance, the fundamental thought of this technique is Principle Component Analysis (PCA), it is straight forward and get a decent execution in little dataset. Be that as it may, it is touchy to enlightenment when the appearances

in the same light can be view as a similar individual. Contrasting with the traditional face acknowledgment techniques, profound learning appears have a superior preferences in enormous dataset, can defeat the assorted circumstances and can separate the ideal confront highlights. It is accounted for that the best execution in Labeled Face in the Wild (LFW) test dataset is around 60% exactness utilizing traditional strategies while getting 99.47% precision utilizing profound learning. Some prominent variables for face acknowledgment innovation not utilized generally today are speed and exactness. Indeed, even profound learning gives a decent execution of exactness, the vast measure of figuring make it hard to use in ordinary PC or implanting gadgets and the speed of acknowledgment is likewise a major issue. Also, taking in a profound system for face acknowledgment require a enormous dataset, it is additionally hard for us in the ordinary research work. For such issues, we have proposed a novel technique for dataset enlarge and a streamlined face acknowledgment framework that can enhance the exactness and speed for face acknowledgment. Whatever remains of this paper is composed as takes after: in segment 2, firstly, we have presented the pipeline of face acknowledgment framework. At that point, we utilize some picture preprocessing strategies prior to the face recognition. In segment 3, we have portrayed the confront acknowledgment framework and the technique for dataset expand. A profound learning calculation demonstrate called Convolutional Neural Network (CNN) have been presented taken after. In segment 4, we give the outcomes for face acknowledgment utilizing CNN calculation.

2. Methodology

2.1. The Face Recognition System

The face acknowledgment framework ought to have the accompanying stages: picture catch, picture reprocessing, face identification and face acknowledgment. As shows in Figure 1, we ought to catch a picture at to start with, then preprocessing this picture for enhancing the nature of the picture and identify the picture whether have a face.
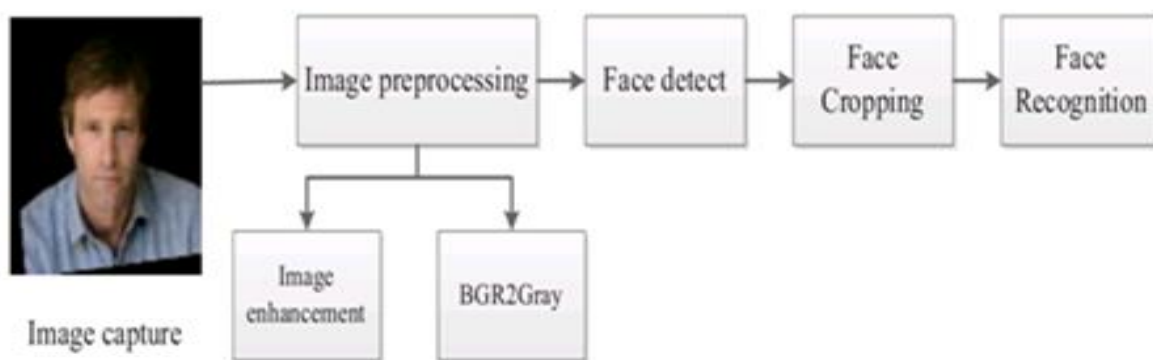

Figure 1: Face Recognition System

In the event that it has identified one face, we will edit it and send it to the profound net to extricate the component vector for acknowledgment. Finally, we will contrast the element vector and the element dataset that was manufactured in the past. The techniques that can calculate the

likeness of two element vectors we regularly utilize have Euclidean separation and angle cosine strategy. In this paper we will utilize the point cosine to calculate the comparability of two component vectors. We will view two faces as a similar individual if the similitude of their element vectors is higher than the edge.

2.2. Image Capture and Preprocessing

The pictures which are caught regularly frequently have numerous arbitrary noises. On the off chance that this picture send to face discovery and acknowledgment frameworks straight forwardly, the outcomes are constantly extremely poor. A sensible preprocessing for pictures will enhance the outcomes exceedingly.

But a novel algorithm for the light impact or countenances, a picture preprocessing calculation for multi poses impact in face recognition and a cross age preprocessing technique for face acknowledgement frame work have proposed to overcome the above problem.

Here, we utilize the quickest face location calculation in light of Haar highlight. As Haar components can just use in the dim space of pictures, we change the picture from RGB space to dark space firstly. Keeping in mind the end goal to make the differentiate proportion more grounded and enhance the nature of the pictures, we pick the least complex picture improvement calculation: the direct dim change calculation.

By and large, if the picture was impacted by an outside environment, we will deteriorate picture quality and the range of dark space was extremely tight. With a specific end goal to develop the scope of dark space, we utilize the direct dim change. Expecting the pixel of unique picture is f(x,y). The pixel estimation of target picture is g(x,y). The dark scope of unique picture is (fmin,fmax) and the objective range is (gmin,gmax) .The social diagram as shows in Figure 2. Utilizing the addition change work appears in equation (1).

$$g(x,y) = \frac{f(x,y) - f_{min}}{f_{max} - f_{min}}(g_{max} - g_{min}) + g_{min}$$

$$= \frac{g_{max} - g_{min}}{f_{max} - f_{min}}f(x,y)$$
$$+ \frac{g_{min}(f_{max} - f_{min}) - f_{min}(g_{max} - g_{min})}{f_{max} - f_{min}} \quad (1)$$

Where (x,y) is the area of pixel. fmax ,fmin is the least and most extreme pixel estimations of unique picture. gmin ,gmax are the base and most extreme pixel estimation of target picture. The picture we utilized ordinarily is 8 bit, so the greatest pixel esteem is 28=256, so the scope of target picture is [0,255]. It implies that gmin=0 and gmax=255. So the equation (1) can be changed to the formal of equation (2). As shows in Figure 2, the nature of picture is enhanced and the scope of histogram is all the more to a great extent.
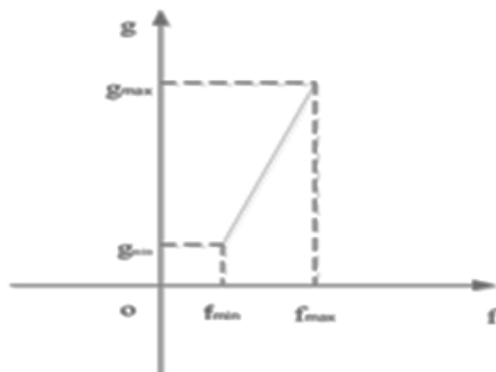
Figure 2: Mapping Relation between f and g

The picture we utilized ordinarily is 8 bit, so the greatest pixel esteem is 28=256, so the scope of target picture is [0,255]. It implies that gmin=0 and gmax=255. So the equation (1) can be changed to the formal of equation (2). As shows in Figure 3, the nature of picture is enhanced and the scope of histogram is all the more to a great extent.

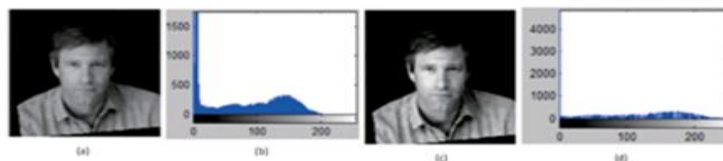$$g(x,y) = \frac{255f(x,y)}{f_{max} - f_{min}} - \frac{255f_{min}}{f_{max} - f_{min}} \qquad (2)$$

Figure 3: The Image Quality That Have Enhanced

2.3. The improvement of Face Detection Algorithm

In this paper, despite everything we utilize the face recognition framework that in light of Haar elements. For the most part, the face indicator prepared by Haar includes dependably have a decent execution for front countenances what's more, bed execution for the appearances that turn points. The Haar highlights we regularly utilized as show as a part of Figure 4. Another issue is that face just hold little zone of a picture. While, It will take quite a while when indicator identify the entire territory of the picture. With a specific end goal to conquer these issues, we propose a new discovery pipeline here. The location pipeline appears in Figure 5. For an info picture, subsequent to preprocessing the first picture, confront finder will recognize the range of the picture, if not identify one face, the picture will turn (±10deg., ±20deg.) and recognize again to beat the revolution confront that can't distinguish. As appears in Figure 6, the face in unique picture is not.

Figure 4: The Haar Features

frontal, through these changes, we can see the locator can distinguish the picture in the first and turn 10 edge, 20 point pictures when can't recognize in the pictures that have pivot - 10 what's more, -20 edges.
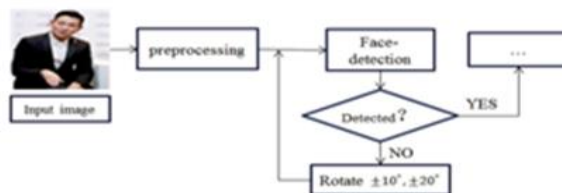

Figure 5: The Path of Face Detection

Figure 6: The Detect Results of the Original and Rotation Images

For an info picture, we preprocess it including change RGB space to dark space and upgrade the quality. At that point, separate the picture skin territory and the locator just finder the skin region that additionally the conceivable face range. Through this operation, the speed of recognition enhanced profoundly. For skin extraction, we ought to change the RGB space to Y Cr Cb space. The change work appears in equation (3). Became very small compare to the original one. The skin extraction sketch shows in Figure 7.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \frac{1}{256}\begin{bmatrix} 65.738 & 129.057 & 25.06 \\ -37.945 & -74.494 & 112.43 \\ 112.439 & -94.154 & -18.28 \end{bmatrix}\begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (3)$$

Through the skin extraction and rotation, the accuracy and speed of face detection have improved synchronously. The discussion of speed and accuracy will show in part IV.

Figure 7: The Skin Extraction Images

2.4. Face Cropping

After face detection, we will crop the face area into the face recognition network for feature extraction. Then, compare with the feature dataset and identify the face. There are two plans that can be picked. The one is just trimming the face zone and resize to the predefined estimate when

another is restrict the eyes as focus and grow to the predefined measure along upward, descending, leftward and rightward bearing. As shows in Figure 8, (a) is the first pictures, (b) is the trimmed pictures that lone have confront territories and (c) is edited pictures utilizing technique.

Figure 8: Cropped Images

2.5. Big Datasets Augment and Deep Face Recognition

Deep face acknowledgment turned out to be exceptionally famous as of late for its advances in pattern acknowledgment and huge information arrangement. One basic condition for profound inclining is huge dataset. Scientists observed that utilizing more information when preparing the acknowledgment model will get a more pleasant execution. Facebook utilizing 4.4M face pictures for training. Furthermore, VGG display likewise have 2.6M pictures for preparing. They are all huge dataset. The pioneer of web like Google, FaceBook possess the huge dataset however they don't contribute them. VGG have proposed a strategy for information gathering however additionally cost huge work for people and need quite a while. Table 1 demonstrates the status of the enormous face dataset. They are all enormous face dataset however few of them share the assets.

TABLE I.   THE BIG FACE DATASET

Num	Name	ids	images
1	LFW	5749	13233
2	WDRef	2995	99773
3	CASIA-WebFace	10575	494414
4	FaceBook	4030	4.4M
5	Google	8M	200M
6	VGG	2622	1635159

2.6. Deep Face Recognition Algorithm

Deep learning is the most prevalent machine learning technique as of late that utilized as a part of face acknowledgment challenge. In this paper, we will utilize the Convolutional Neural Network (CNN) for face acknowledgment and reference to the VGG arrange. We utilizing this VGG-NET to separate the face highlight and contrast with the exist dataset. The primary layers of CNN have convolutional layers (CL), pooling layers and full association layers. In CL, we can

respect this layer is to concentrate highlights. After convolutional, we regularly make a dynamic operation. The convolutional operation appears in Figure 9. Convolutional operation is a numerous to numerous mapping relationship.
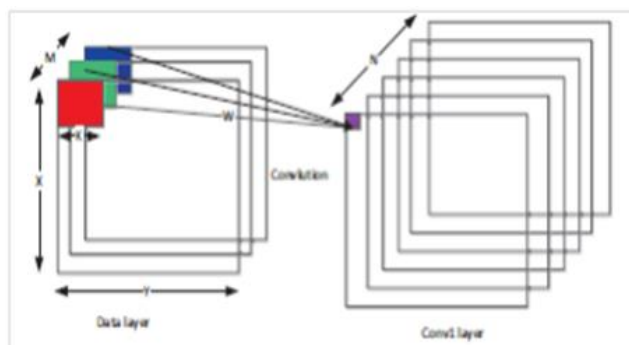

Figure 9: Convolutional Operation

We can separate the key components through convolution for ordinarily. The equation of convolution is (14).

$$y_l(n; x, y) = y_{l-1}(m, x + p, y + q) * W(m, n; p, q) \quad (14)$$

Where, l speak to the lth layer, l-1 speak to the (l-1)th layer and W is the parameters. One issue is that the convolution brought about numerous parameters and over fitting, we can diminish the parameter through pooling operation. We regularly utilize max-pooling operation in the system. Max-pooling takes the most extreme estimation of one territory. Along these lines, the measure of highlight maps will diminish. The equation appears in (15). Different routes for diminish the parameters and over fitting like commencement, inserting, and dropout.

$$pooling_{max}(R_k) = \max_{i \in R_k} a_i \quad (15)$$

Where, ai is the estimation of region Rk.

3. Test Result

This paper, have three experiments. The first one is checking the improved algorithm of face detection. The second one is face feature extraction using deep network based on VGG-NET. The third one is using the VGG-NET for fine tune on our own dataset.

3.1. Results Comes Out

An enhancing face discovery calculation that have depicted some time recently. Here, we get the speed and exactness of given calculation. The test dataset have 200 countenances that chose in LFW dataset and resize them to 256*256. As shows in Table 2, we can see the recognize exactness was enhanced exceedingly when improve the nature of pictures. At the point when pivot the picture for further face recognition, the distinguish exactness can likewise enhance,

though the time costs a great deal. Include the face skin extraction contribute little to the identify precision furthermore squander additional time.

TABLE. II   THE RESULTS OF FACE DETECTION

NUM	Method	Speed	Acc
1	Traditional	256ms	0.8023
2	Traditional+iamge enhancement	245ms	0.8552
3	Traditional+image enhancement+image rotation	500ms	0.8878
4	Traditional+image enhancement+image rotation+skin extraction	720ms	0.8858

In the second analysis, we extricate the face include in light of the VGG-Net. We select two pictures that have a place with same individual and one picture that have a place with other. We measurable the highlight vector in FC7 layer. As shows in Fig. 10, we can see the measurable figures of second and third pictures are fundamentally the same as at the point when are distinctive to the primary picture. In this paper, we utilize the edge cosine esteem to gauge the closeness of two vectors. Accept x and y are tow vectors, utilizing edge cosine esteem work as shows in (16). The greater the esteem is, the more similitude the two vectors are.

$$\cos(x,y) = \frac{(x,y)}{\|x\| \cdot \|y\|} = \frac{\sum_{i=1}^{n} x_i \cdot y_i}{(\sum_{i=1}^{n} x_i^2 \cdot \sum_{i=1}^{n} y_i^2)^{\frac{1}{2}}} \quad (16)$$

Utilizing equation (16), we ascertain the similitude of these three pictures appears in Figure 10. The similitude between first picture and second picture is 0.4256, the first and third is 0.3652.
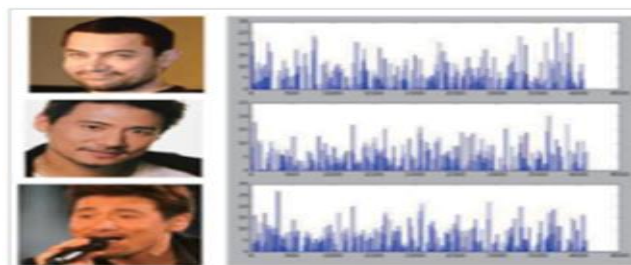
Figure 10: The Feature Vector in FC7 Layer

At the point when the second and third is 0.7256. Also, we have chosen 900 sets pictures in the LFW dataset, the initial 300 sets are certain examples (have a place with same individual), the center 300 sets are pessimistic examples (distinctive people) and the last 300 sets are sure among them. As shows in Figure 11, we have measurable the similitudes of each match.
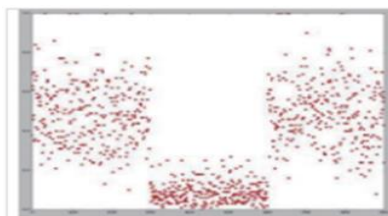
Figure 11: The Statistical of Every Pair Faces

We can see the grouping characters is great, the likenesses of same people (constructive examples) are higher at the point when the likenesses of various people (pessimistic examples) is lower. The exactness touched base at 0.97 on the off chance that we set the limit as 0.25.
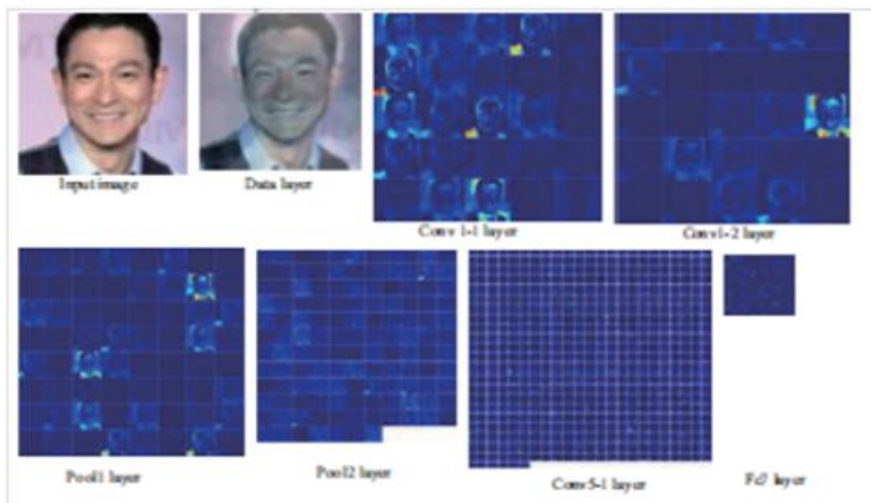

Figure 12: The Features in Some Layer

But, we can see the component in each layer of the net.

Here, we take the info information, information layer, conv1-1 layer, conv1-2 layer, pool 1 layer, pool 2 layer, conv 5-1 layer and FC7 layer for example. The features show in Figure 12.

 In the third experiment, we fine-tune the VGG-NET in our own dataset. We choose four famous stars and select 50 images for each of them on the internet. We can get 200 images for each star after using the dataset augment algorithm we have described before. Here, we choose 150 images per person for training and 50 images for testing. The initial learning rate we set as 0.001, and decrease the learning rate 10 times after 100 iterations. As shows in Fig. 13, the loss value decrease frequently and get the accuracy about 0.9.
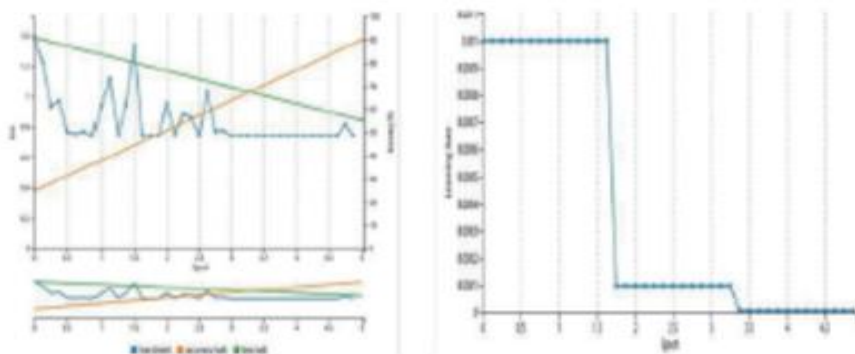

Figure 13: The Loss Curve and Learning Rate Curve

Finally, we have drawn the Receiver Operating Characteristic (ROC) curve. Through the ROC curve, we can watch that the certainty level of this model is high. ROC bend appears in Fig. 14.
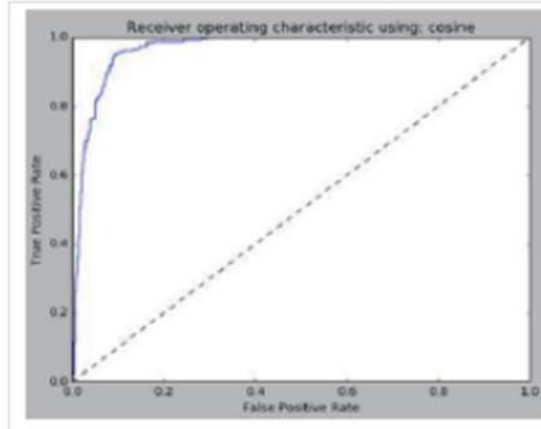
Figure 14: ROC Curve

4. Conclusion

This paper proposed a cross breed preprocess of pictures. That enhanced the exactness of face recognition and acknowledgment to some degree. The technique to enormous dataset enlarge also, utilizing VGG-Net for calibrate gives a decent execution. Be that as it may, a perfect face acknowledgment framework ought to get a higher precision and speed. A lot of calculation make the framework can't have any significant bearing on the inserting frameworks. In addition, a great framework ought to hearty to the brightening, multi-postures, expression and cross-age. We trust that, with the improvement of equipment and enhancing calculation, shutting the crevice between the human and machine is not a fantasy.

References

[1] Kwak K C, Pedrycz W. Face recognition using a fuzzy fisherface classifier [J]. Pattern Recognition, 2005, 38(10): 1717-1732.
[2] Wang X, Han T X, Yan S. An HOG-LBP human detector with partial occlusion handling [C]. Computer Vision, 2009 IEEE 12th International Conference on. IEEE, 2009: 32-39.
[3] Sun Y, Wang X, Tang X. Deeply learned face representations are sparse, selective, and robust [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 2892-2900.
[4] Xu Y, Zhang Z, Lu G, et al. Approximately symmetrical face images for image preprocessing in face recognition and sparse representation based classification[J]. Pattern Recognition, 2016, 54: 68-82.
[5] AbdAlmageed W, Wua Y, Rawlsa S, et al. Face recognition using deep multi-pose representations [J]. arXiv preprint arXiv:1603.07388, 2016.
[6] Taigman Y, Yang M, Ranzato M A, et al. Deepface: Closing the gap to human-level performance in face verification [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014: 1701-1708.
[7] Parkhi O M, Vedaldi A, Zisserman A. Deep face recognition [J]. Proceedings of the British Machine Vision, 2015, 1(3): 6

*Corresponding author.
E-mail address: prashantkushwah211@ gmail.com