Hacettepe University

Computer Engineering Department

BBM479 End of Term Development Report

Project Details

Title	Class Attendance With Face Recognition
Supervisor	Nazlı İkizler Cinbiş

Group Members

	Full Name	**Student ID**
1	Enes Taha Elmaci	21990872
2	Pelin Fildiş	21827412
3	Abdulkadir Alaca	21827019
4		

Current State (/ 40 Points)

Explain the current state of the project. By giving references to the plan proposed at the beginning of the project, explain what is achieved so far. Provide details of what has been done by explaining the technology and methodology used. Is the current state of the project inline with the plan? If you are behind the schedule, explain in detail the reasons.

In our project, we aimed to create a system that uses face recognition techniques with artificial intelligence to take attendance in a classroom. In order to achieve the goal of this system, we first need to extract the faces of students from a photograph taken in the classroom. Then, we need to give the extracted faces to a trained model that can identify the individuals. To achieve this goal, we first conducted a literature review and consulted with people who had knowledge about this topic, as well as looked at previous studies on this subject. At the same time as the literature review, we began creating our data set. We asked ten of our friends to take photographs from different angles and used these to create the data set.

As a result of our research, we first used a pre-trained neural network called FaceNet[1] to detect faces in a photograph. In our trials, we found that it was able to successfully select faces from group photographs.

After completing this step, it was time to identify the individuals. We first began training a ResNet34[2] structure using our data set, but this method was not suitable for our problem. The main problem in our project is to be able to recognize faces with very few (1 or 2) photographs for each person. We would need a very large data set and a lot of time to train ResNet34 to solve the face recognition problem. We spent two weeks trying to get results from this neural network and improving the results we had. But we could not reach our goals about accuracy. Upon consulting with our supervisor, we were recommended to focus on Siamese neural network structures and metric learning methods.

After this advice, we completed our research on metric learning and Siamese networks. Based on the literature review, we thought that one-shot learning might be a solution to our problem, so we focused on papers[4] and studies related to this topic. We began testing these methods and achieved very successful results in a short time. We used the VGG16[3] structure as a Siamese twin. We used the pretrained weights of the VGG16 trained on the data set created for face recognition. We obtained these weights from Kaggle. After loading the weights into the model, we modified the last layer of the model. Instead of using it as a classification layer, we used it as an embedding layer. Then, we fed the results from the embedding layer for two different photographs given to the model to the "Cosine Similarity" function in order to determine the similarity between the two faces.

Although there are areas where our model needs improvement, it is currently able to recognize faces with great accuracy. So, our project is almost inline with the plan. In the next stages of our project, we plan to improve our model and try some new techniques and then move on to the web (application) part as soon as possible.

Continuous Learning (/ 25 Points)

You have been working on the problem for almost 3 months. In these 3 months, what did you learn about this problem that you didn't know at the beginning? Did this new knowledge change your perspective on the problem? What else do you need to learn in the future?

At the beginning, we started with finding faces in the photo. We used dlib and MTCNN. But MTCNN is better than Dlib. We thought we could train a model with 10 or 20 pictures. With this approach, we found papers that explain ResNet34. After that, we started to implement that architecture. But we get bad results because our dataset is not enough to train a good model and the ResNet34 structure is not suitable for our problem.

After the failure, we decided to change our search keywords and main focus neural network model. We started to search papers about one shot image recognition, we did not only use face recognition keyword but alo face identification etc.. While we were searching, we found a paper about the VGG-16 model. We used that model's weights which are trained with 15 million labeled images. We started to implement that architecture. We defined a cosine similarity function to identify faces. We gave 10 people's photos to train our model. Then we gave some photos of these people in those photos. Our model identifies them if the cosine similarity is below the threshold that we gave(we tried with 0.30). We test our model with photos taken between one and seven years and with long hair/short hair, with beard or shaved face. But it is still not good enough. We need to make our model better.

We need a better face detector because sometimes faces that are inclined cannot be detected as faces and even if detected sometimes our model could not identify the person.We found YOLO has a good accuracy to detect faces. We are still searching for improving our model.

During this period, we learned how to work as a team. In the following period, we learned that we should have regular meetings every week. We also learned that we need to do our research more comprehensively in order not to waste time from now on.

Risk Assessment and Management (/ 20 Points)

Try to identify the potential risks in the rest of the project. Were you able to identify these risks at the beginning of the project, or did you recently recognize them? What are your proposed solutions to each risk item? Are there any risks that will require a significant change in the project? If so, explain how this will affect the end results (also, outline your proposed revisions in the next section).

There are some potential risks in our project. Most of the risks could occur after the release. In the rest of the project we may have trouble with gpu capacity while training the model. We do not have high performance computers so our machines are not good for training. As a solution, we are trying to use Google CoLab to train model. Our project will be used in real life and changes always happen in real life. Face of students can be changed. For example, make up, beard, glasses or some differences on faces. To handle this risk, we are trying to use static situations on the face. Also to train a model, there needs to be a lot of photos in the dataset. It is hard to have a lot of photos of each student, so we decided to use one shot learning, Siamese Network and metric learning which can train with only one photo for each face.

In recent history we had a pandemic. There may be a pandemic again and if the people had to use masks on their face again, our model may not work well. We do not have a solution for this risk now, because this risk can cause big changes in our project like change of model or camera system. Another risk is the size of the classes, neural networks may not be able to detect the faces of students who sit on the backside of class. We can improve the detection ability of the network by taking more than one snapshots with different zoom ratios and different angles.

In testing, we got permission from our friends to use their photos but if this project will be used in real life, there can be some problems about ‘KVKK’. We cannot solve this problem, the institutions who want to use this project need to solve it with their own students.

Revisions (/ 15 Points)

If you feel like you need to revise your earlier plan, suggest your changes here. These changes may include changes in outcomes of the project, changes in milestones, changes in calendar, changes in workload distribution, etc. If you do not present any revisions here, at the end of the project you will be responsible for all the proposed outcomes in the Project Proposal.

As we continue the project in accordance with our first plan, there were some minor setbacks as we mentioned in the previous sections. We lost a couple of weeks trying to implement ResNet34, which was the wrong approach for our project, and trying to improve the results we obtained from this method. However, thanks to the dedicated work of the three of us, we were able to make up for this deficiency. Due to the previous setbacks, maybe it took us a couple of extra days to start the backend and frontend parts.

In terms of task division, at the beginning of the project, we thought we would equally take care of the machine learning and model-related parts. However, as the process progressed, Enes took care of the face detection part, Kadir and Pelin took care of the face recognition/identification part, and equal attention was given to one-shot learning and metric learning towards the end. During the creation of the dataset, we all worked together and took photos of our friends with their permission. In later stages, Pelin will focus on improving the model, Kadir will focus on the backend, and Enes will focus on the frontend.

If the efforts to improve the model's accuracy and fine-tune part take longer than expected, it will not affect the web part of the project because we planned to carry out the model improvement/testing, backend and frontend stages in parallel. However, if the AI phase takes longer than expected, there may be delays in integrating and testing the entire system. However, we are working hard as a group before starting the web part to avoid this.

We have learned many lessons from the research and work we have done this term, and from now on, we will be careful not to make the mistakes we made this term while developing our project. For example, in the next stages, we will start to develop a method/technique after we research better before trying it, get ideas from our supervisor more often and question the suitability of the method/technique we find for the problem.

References

[1]https://www.researchgate.net/publication/304552788_FaceNet_A_unified_embedding_for_face_recognition_and_clustering

[2] https://paperswithcode.com/paper/deep-residual-learning-for-image-recognition

[3] https://paperswithcode.com/paper/very-deep-convolutional-networks-for-large

[4] https://www.cs.cmu.edu/~rsalakhu/papers/oneshot1.pdf