Intelligent Photo Gallery

Detailed Interim Report

Kwan Hoi Shun (3035280972)

Supervised by Dr. Anthony Tam

January 2018
Abstract

People nowadays take a lot of photos and are in need of a tool for organizing their collection of photos. Existing photo gallery application such as Google Photos uses face recognition and cloud storage to improve the photo management experience. However, the cloud storage poses privacy issues and using the local machine for face recognition can be slow.

This project aims to build a photo gallery application to provide an alternative and improve user’s photo management experience. It also aims to find and efficient face recognition model to be incorporated into the application.

A working prototype has been built in the current stage and the basic functionalities has been implemented into the prototype. Currently, we are working on the face recognition feature. Once the testing and integration of the face recognition model is finished, we will move on to the development of the advanced features of the application.
Acknowledgment
We would like to thank our supervisor, Dr. Anthony Tam for providing advices and insight during the development of the application.
# Table of Contents

**List of Figures** ......................................................................................................................... 5  
**List of Tables** .......................................................................................................................... 5

**Chapter 1.0 Introduction** ........................................................................................................ 6  
1.1 Background .......................................................................................................................... 6
1.2 Existing Solutions .................................................................................................................. 6  
1.2.1 Google Photos .................................................................................................................. 6  
1.2.2 Facebook ......................................................................................................................... 7
1.3 Project Objective ..................................................................................................................... 7
1.4 Outline .................................................................................................................................... 7

**Chapter 2.0 Project Scope** ...................................................................................................... 8  
2.1 Major Features ....................................................................................................................... 8
2.2 Feature Priority List .............................................................................................................. 9
2.3 Project Deliverable ................................................................................................................ 10

**Chapter 3.0 Methodology for Face Recognition** .................................................................... 11  
3.1 Face Detection Stage .......................................................................................................... 11
3.2 Feature Extraction Stage ...................................................................................................... 11  
3.2.1 Face Recognition by extracting facial features ................................................................. 12  
3.2.2 Face Recognition by learning facial features ................................................................. 12
3.3 Classification Stage .............................................................................................................. 12
3.4 Tests for the facial recognition algorithms .......................................................................... 13  
3.4.1 Test for Feature Extraction algorithms ......................................................................... 13  
3.4.2 Test for Classification algorithms ................................................................................ 13

**Chapter 4.0 Development Process** ...................................................................................... 14  
4.1 Design and Planning Phase ................................................................................................. 14
4.2 Implementation Phase ......................................................................................................... 14  
4.2.1 Basic Features Implementation Stage ........................................................................... 14  
4.2.2 Face Recognition Stage .................................................................................................. 14  
4.2.3 Advanced Features Implementation Stage .................................................................... 14
4.3 Optimization and Testing Phase ......................................................................................... 15
4.4 Development Tools and Technologies .............................................................................. 15  
4.4.1 Application Language: Python ...................................................................................... 15  
4.4.2 GUI Framework: Kivy ..................................................................................................... 15  
4.4.3 Image Processing Library: OpenCV and Pillow ............................................................ 15  
4.4.4 Database Software: Sqlite3 ........................................................................................... 15

**Chapter 5.0 Progress** ........................................................................................................... 16  
5.1 Current Application Development ...................................................................................... 16  
5.1.1 Import Photos ................................................................................................................. 16  
5.1.2 Search Photos .................................................................................................................. 16  
5.1.3 View Photo ...................................................................................................................... 17  
5.1.4 View Album ..................................................................................................................... 18
Currently, we are working on the face recognition stage of the implementation phase. The face recognition approach has been decided and we are working on building the face recognition models for comparisons. The face detection stage has been built using the functions from openCV, and we are now testing the feature extraction algorithms with the labeled faces in the wild dataset.

5.3 Future Work

5.3.1 Future Application Development

5.3.2 Optimizing and Testing the Application

5.3 Schedule

Chapter 6.0 Limitations and Difficulties

6.1 Limitation – Training feature extraction algorithms

6.2 Difficulty – Kivy GUI Framework

Chapter 7.0 Conclusion

References

List of Figures

Figure 1: The flow of the Facial recognition approaches
Figure 3: The photo import screen
Figure 4: The search photo screen
Figure 5: The image view screen
Figure 6: Create album popup
Figure 7: Album Screen
Figure 8: View album screen

List of Tables

Table 1: The priority of each of the feature. The priority are ranked in the scale of 1-5, with 1 being the most important and 5 being least important. Features with priority larger or equal to 3 are considered core features for the application and features of lower priority are tentative features.
Table 2: The project schedule
Chapter 1.0 Introduction

1.1 Background

Nowadays, people take a lot of photos with their smartphones. Many people have accumulated a large collection of photos without proper management in the collection. A good photo gallery application can allow them to organize their built up collection of photos and locate photos in their collection with ease through the use of tags, albums and other photo management functionalities.

Face recognition and cloud storage has been implemented in some of the existing photo gallery application. Face recognition allows user to search for specific photos more easily by adding more information about the photos while cloud storage make users to access photos more efficiently. However, the use of cloud storage also raises concerns over the privacy of the photos and the speed to access the photos is limited by the internet bandwidth.

1.2 Existing Solutions

Currently, there are many social network application that allows user to share photos with others, however social network applications such as Facebook and Instagram are not alternatives to a photo gallery application as they compresses the photos resulting in a lower image quality. The intelligent photo gallery can be better compared to photo gallery applications such as Google Photos and Photos (Mac).

1.2.1 Google Photos

Google Photos is a popular photo gallery application make use of cloud storage to allow users to access their photos easily, it also have a quick face recognition feature. However user have to face a choice between limited storage and reduced photo quality if they use the application for free. Moreover, the application requires the use of cloud storage which raises concerns on the privacy as you loses control over the photos once they are uploaded to the cloud storage.
1.2.2 Facebook
Another popular choice for people to store their photo is Facebook. Facebook also provides face recognition feature along with other photo management features such as albums, it also allows people share their own photo easily. However, Facebook uses compression to store the photos, hence the photos are stored with reduced quality. Facebook is also known to data mine user information, therefore it also raises concerns whether it is safe to store photos on Facebook where Facebook may potentially analyze the uploaded photos for suggesting advertisements.

1.3 Project Objective
This project aims to build a photo gallery application that aims to protect user’s privacy by storing them in a private storage device and enhance the user’s photos management experience by introducing more efficient face recognition algorithms and innovative photo management features.

Different face recognition algorithms will be tested in this project and the best model, judged by its efficiency and accuracy amongst the models, will be incorporated into the application.

For storage options, the intelligent photo gallery will allow the user to choose storing their photos in their local device or a network attached storage (NAS) device. The NAS device is essentially hard disks connected to the local network. This allow users to access the photos without bandwidth limitation and multiple users can access the photos within the local network.

1.4 Outline
This report will first discuss the background and objective of this project in Chapter 1. It will then carry on with the project scope, major features and the priority the features in Chapter 2. Afterwards, it will discuss the methodology of the face recognition that will be used, the development process and the technology and tools involved in this project in Chapter 3 and 4. Then, it will review the progress of the progress and mention the limitation and difficulties encountered in the development of the application in Chapter 5 and 6. Finally it will conclude in Chapter 7.
Chapter 2.0 Project Scope

This project will build a desktop photo gallery application that aims to enhance user’s photo management experience. The application will include basic photo gallery functionalities (i.e. import photos, search, tagging), more advanced photo management features (i.e. duplicate photo search, automatic photo filters) and an efficient face recognition model.

2.1 Major Features

- **Face recognition for recognizing faces**
  The application will have an efficient face recognition algorithm that can identify faces in photos accurately. The face recognition process will be carried out during the importation of the photos, and users will be asked to provide names of the faces where it has been identified multiple times.

- **Import photos and search photos**
  The application will allow users to import and export photo and have a quick search functions that can help users to locate their photos quickly.

- **Duplicate search for searching duplicate photos**
  A duplicate search function will be implemented into the application to help users to identify duplicated photos.

- **Add and remove tags and albums to photos**
  User will be able to use tags and albums to organize photos in groups. Tags and albums are used to organized photo into different groups and users can use tags and albums for searching. A key difference between tags and albums is photos can have several different tags, while it can only belong to a single album, hence users should use albums to group photos by the most important characteristic (i.e. important events) while tags should be used to provide less significant characteristics (i.e. objects in the photo, emotion)

- **Apply actions to selected photos**
  Users will be able to apply the same action (e.g. adding tags) to a selection of photos.
- **Auto filters for importing photos**
  The auto filters are very similar to email filters. Users will be able to specify criteria and action to be done such that specified action will be applied automatically to all newly imported photos satisfying the criteria.

- **Storing the photos in local storage and in the NAS storage device**
  Users will be able to configure the storage options (local storage and NAS storage) of the application. More storage options including NAS storage with personal cloud configuration will be added if time is allowed.

- **User friendly Graphical User Interface (GUI)**
  The GUI of the application will be intuitive and users will be able to locate and use the functions of the application with ease.

### 2.2 Feature Priority List
Since there are quite a lot of features to be implemented in this project, a priority has been assigned to each of the features to indicate its importance (see Table 1). The project will then implement the features in decreasing order of priority to ensure that all important features will be included in the final deliverable.
Table 1: The priority of each of the feature. The priority are ranked in the scale of 1-5, with 1 being the most important and 5 being least important. Features with priority larger or equal to 3 are considered core features for the application and features of lower priority are tentative features.

2.3 Project Deliverable

This project will produce a desktop photo gallery application that will include all the feature of priority 3 and higher. The application will be compatible on both Windows and Mac.
Chapter 3.0 Methodology for Face Recognition

This project will compare and evaluate two face recognition approaches on their accuracy and efficiency. The approaches can be divided into 3 stages, namely face detection stage, feature extraction stage and classification stage. The two approach will use the same face detection and classification procedures with difference in the feature extraction stage (see Figure 1).

3.1 Face Detection Stage

Facial regions will be located and cropped during the face detection stage. This stage will be implemented using functions from the OpenCV library as face detection is a well-established problem and the corresponding functions from Open CV are highly optimized. The local binary pattern classifier is chosen as the algorithm over the HAAR classifier as it is much more efficient. The reduced accuracy of LBP is not significant as when people take photos, they position their face towards the camera, hence the photo taken will have faces that are easy to identify.

3.2 Feature Extraction Stage

In the feature extraction stage, the facial regions will be reduced to a number of features. The two different approaches will use machine learning to extract different features. The models will be trained in the development phase and the final application will used the trained model for extracting feature. The details of the two approaches are given in the following section. (see Section 3.2.1 and Section 3.2.2)
3.2.1 Face Recognition by extracting facial features

This approach uses a machine learning based face landmark estimation algorithm to locate the facial landmarks (e.g. eye, nose, mouth, etc.). The machine learning algorithm learns to find the specific facial landmarks and use the ratio between different facial landmarks(e.g. ratio between the eye-nose distance and the eye-chin distance) as features for classification. This approach is invented by Ramesha, Raja, Venugopal and Patnaik[1], however in the original approach the facial features are located using cranny edge detection where this project uses a landmark estimation algorithm to locate the facial landmarks.

In addition to the ratio between different facial landmark features, other facial features such as the eye colour and face colour will also be extracted from the facial features, this is to provide more information for the classification stage. The eye colour can be determined as we can use the location of the eye determined by the facial landmark algorithm.

3.2.2 Face Recognition by learning facial features

Different to the previous approach, the feature points extracted in this approach will be learned by a machine learning algorithm. This approach is very similar to the approach used by facenet invented by Schroff, F., Kalenichenko, D., & Philbin, J. [2] The face recognition algorithm will find the best 128 features to distinguish different persons. A deep convolutional neural network is be used to select the facial points, and in each training iteration, a triplet of photos is used, two photos from the same person and a third photo from a different person. The network will then be tweaked such that the feature points extracted from the same person will be more similar and vice versa.

3.3 Classification Stage

In the classification stage, the features extracted in the feature extraction stage are be used to classify the face into different persons. The extracted features will then be used to classify the person of the face. Different classification algorithms including support vector machines(SVM), k nearest neighbours and clustering will be used to determine the best classification algorithm for this stage.
3.4 Tests for the facial recognition algorithms

3.4.1 Test for Feature Extraction algorithms
For this test the Labeled Faces in the Wild dataset is used to determine which algorithm is best at differentiating faces belongs to different person. [3] This dataset is widely used as a test for face recognition algorithms hence is chosen for the test for the feature extraction algorithms. The dataset consist of many pairs of faces and the algorithms needs to determine whether the pair of faces belongs to the same person or not.

3.4.2 Test for Classification algorithms
As the Labeled Faces in the Wild focuses on how good the algorithm is determining faces belongs to different persons, we need another test to determine how good the classification algorithms are at learning and classifying faces belongs to different. For this test the VGGFace2 dataset is chosen as it has a large amount of face images that belongs to different persons, with an average of 362 faces per subject.[4] In this test, faces belongs to each subject will be split into a train set and test set, and the classification algorithms will need to determine the person of each of the faces in the test set based on the training using the train set.
Chapter 4.0 Development Process

The development process of this project consist of 3 phases, namely design and planning phase, implementation phase and optimization and testing phase. The following sections will provide details about the work to be done in each of the phases.

4.1 Design and Planning Phase

During the Design and Planning phase, the features and scope of the project are finalized. The technologies that will be involved in the project are also determined at this stage. This includes the database, GUI framework, etc.

4.2 Implementation Phase

During this phase, the features will be implemented into the application. The implementation phase is divided into 3 smaller stages to allow the application to be built iteratively. Each stage in the implementation phase will implement a partition of the features into the application.

4.2.1 Basic Features Implementation Stage

In this stage, the basic features will be implemented into the application. This includes importing and search photos and implementation of tags and albums. After this stage a simple working photo gallery application prototype will be developed.

4.2.2 Face Recognition Stage

In this stage, the models for the 2 face recognition approaches will be built, trained and tested for their accuracy and efficiency. The better model will be incorporated into the application at the end of this stage.

4.2.3 Advanced Features Implementation Stage

In this stage, the advanced features such as photo filters and duplicate photo search will be implemented into the application. After this stage, all the core functionalities will be implemented into the photo gallery application. If time is allowed, tentative features will also in this stage.
4.3 Optimization and Testing Phase
In this phase, the application will be debugged and optimized for both performance and user experience, this includes improving the graphical user interface(GUI) of the application.

4.4 Development Tools and Technologies

4.4.1 Application Language: Python
- Python is chosen as the application language as the development team is more familiarized with Python. It is also chosen as it have a large collection of packages for dealing with different problems especially in deep learning, as this project will use machine learning algorithms for face recognition.

4.4.2 GUI Framework: Kivy
- Kivy is used as it is a popular GUI framework and is well documented. It is a free GUI framework and the GUI framework is more modern than the Python inbuilt GUI framework tkinter. Kivy is also a cross-platform GUI framework that gives a universal GUI appearance for both Windows and Mac, hence can save development time for tweaking the GUI twice for both platform.

4.4.3 Image Processing Library: OpenCV and Pillow
- OpenCV and Pillow are be used as the Image Processing library as both library are well-used and well-documented. The libraries are optimized for performance that allows the application to have an efficient image processing operation in the face recognition stage.

4.4.4 Database Software: Sqlite3
- Sqlite3 is chosen over MySQL as the database software as the intelligent photo gallery is a desktop application without any network server. Hence sqlite3 will be sufficient as the database and is simpler to set up than MySQL. The database file of Sqlite3 is also saved as 1 file and can easily transport to and out of a NAS device if the user chooses it as the storage option.
Chapter 5.0 Progress

5.1 Current Application Development

The basic functionalities of the photo gallery application has been implemented in the Basic functionalities implementation stage. This includes the implementation of the import photo, search photo, view photo features and the creation and deletion of tags and albums.

5.1.1 Import Photos

Users can upload photos into the prototype by dragging photos into the upload page of the application prototype. A preview of the photos dragged will be provided in the left side of the page and user can specify the album and tags to be associated with the uploaded photos and confirm the import action by clicking the ‘import’ button. (See Figure 3)

![Figure 2: The photo import screen](image)

5.1.2 Search Photos

The search page of the prototype provide a grid of matching photos to the user. (see Figure 4) The user can specify the keywords to be matched in the search bar and search for desired photos. User can also add ‘t:’ before a keyword to specify photos to have the exact tag name as the keyword and add ’a:’ to specify for photos of the specific album.

For example, to search for photos with tags of name ‘abc’ and album of ‘123’, the user may type in ‘t:abc a:123’ as the search phrase.
5.1.3 View Photo

By clicking on any photo in the search grid. The user can access the detailed image page of the photo. (See Figure 5) Users can also edit details about the image by adding tags, removing tags and changing the album of the photo. The user can also search for photos with any of the tags that is associated with the current photo by clicking on any tags in the bottom of the view image screen.
5.1.4 View Album

Album related activities are performed under the ‘Albums’ screen which can be accessed in the ‘Albums’ button on the left side of the application. User can view all the existing albums in the Albums screen and search for albums with matching name using the search bar. (see figure 7) To create new album, users can click on the ‘Create New Album’ button on the top right of the Album screen and specify the new album name and description in the popup. (see figure 6) Users can also view the photos and description of any of the existing album by clicking on any album in the album screen. (see figure 8)

Figure 5: Create album popup

Figure 6: Album Screen

Figure 7: View album screen
5.2 Current Work
Currently, we are working on the face recognition stage of the implementation phase. The face recognition approach has been decided and we are working on building the face recognition models for comparisons. The face detection stage has been built using the functions from openCV, and we are now testing the feature extraction algorithms with the labeled faces in the wild dataset.

5.3 Future Work
5.3.1 Future Application Development
After testing the feature extraction algorithms we will move on to testing the different classification algorithms. In the meantime of training the classification algorithm for testing, we will start the development of the advanced features which includes the implementation for the photo filters and the duplicate search.

5.3.2 Optimizing and Testing the Application
After developing the core features of the application, we will perform a full test on all of the features of the application and debug the application. In the optimization and testing phase, we will also optimized the application for performance and user experience, this includes reducing the loading time of the different screens and improving the efficiency of the image search feature. We will also improve the GUI such that different user operation will be more user friendly and implement a settings panel for user customizations.
5.3 Schedule

The project will be developed according to the below schedule (see table 2).

<table>
<thead>
<tr>
<th>Stage / Phase</th>
<th>Subtask</th>
<th>Expected time</th>
<th>Expected finish date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic Features Implementation Stage (Priority 1 features)</td>
<td>Design the use case and logic in the basic functionalities</td>
<td>2 weeks</td>
<td>Oct 16</td>
</tr>
<tr>
<td></td>
<td>Implementing the import, view and search function</td>
<td>2 weeks</td>
<td>Oct 30</td>
</tr>
<tr>
<td></td>
<td>Implementing the tags and albums</td>
<td>2 weeks</td>
<td>Nov 13</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>1 week</td>
<td>Nov 20</td>
</tr>
<tr>
<td>Face Recognition Stage</td>
<td>Figuring out how to implement the face recognition algorithms</td>
<td>2 weeks</td>
<td>Dec 4</td>
</tr>
<tr>
<td></td>
<td>Developing the face recognition prototypes and training and testing the models</td>
<td>8 weeks</td>
<td>Jan 29</td>
</tr>
<tr>
<td></td>
<td>Integrating the trained models into the application</td>
<td>1 weeks</td>
<td>Feb 5</td>
</tr>
<tr>
<td>Advanced Features Implementation Stage (Priority 2 features) (Priority 3 features)</td>
<td>Design the use case and logic in the more complex functionalities</td>
<td>1 weeks</td>
<td>Feb 12</td>
</tr>
<tr>
<td></td>
<td>Implement filters</td>
<td>2 weeks</td>
<td>Feb 26</td>
</tr>
<tr>
<td></td>
<td>Implement duplicates search</td>
<td>2 week</td>
<td>Mar 12</td>
</tr>
<tr>
<td></td>
<td>Implementing the NAS storage option and testing</td>
<td>1 week</td>
<td>Mar 19</td>
</tr>
<tr>
<td>Testing and Optimization Stage</td>
<td>Testing and Optimization</td>
<td>5 weeks</td>
<td>Apr 27</td>
</tr>
<tr>
<td>[Extra] Priority 3+ features</td>
<td>If time is available</td>
<td>If time is available</td>
<td></td>
</tr>
</tbody>
</table>

*Table 2: The project schedule*
Chapter 6.0 Limitations and Difficulties

During the development of the application, we have encountered several limitations and difficulties, below are the significant limitations and difficulties we have encountered during the development process

6.1 Limitation – Training feature extraction algorithms

Since the feature extraction algorithms require a lot of computation resources, the facial landmark approach requires the training of the landmark estimation algorithm and the learning of the facial features in the other approach also requires the training of a convolutional neural network. As we do not have the computation resource and time for training the neural networks for millions of photos, we have used trained models for the algorithms instead. However, since the trained model for the algorithms are come from different sources, the training faces for the models are also different, and may result in an advantage for the model that have been trained with more faces.

6.2 Difficulty – Kivy GUI Framework

As the Kivy GUI framework is developed mainly for mobile devices, it lacks some of the functionalities when used for developing desktop application. An example is the lack of event handler for the hover event and functions for showing tooltips. The difficulties faced for developing desktop application with Kivy has been solved by searching for solutions in the community as solutions for the mentioned problems has been shared in the community and the solutions has been implemented into the current application.
Chapter 7.0 Conclusion

This project aims to build a photo gallery application to enhance user’s photo management experience and to find an efficient face recognition algorithm to use in the application.

The progress of the project has been satisfactory as the development of the application is on schedule, the implementation of the basic functionalities has been finished and currently we are working on testing different facial recognition algorithms. Once the testing is finished, we will integrate the best model into the application and start the development of the advanced features.

Once the development of all the important features of the application is finished, we will move on to debugging and optimizing the application for performance the user experience. At the end of this project, we hope to provide an alternative to existing photo management applications that protects user’s privacy and provide a great user experience in managing photo collections.
References


