

**CS 491  
Senior Design Project  
2021 Fall**



**Project Specifications Report  
PolliVidis**

Ömer Ünlüsoy - 21702136  
Elif Gamze Güliter - 21802870  
İrem Tekin - 21803267  
Ece Ünal - 21703149  
Umut Ada Yürüten - 21802410

**Supervisor:** Ercüment Çiçek  
**Jury Members:** Shervin Arashloo and Hamdi Dibeklioglu

This report is submitted to the Department of Computer Engineering of Bilkent University in partial fulfillment of the requirements of the Senior Design Project course CS491/2.

# Table of Contents

- 1. Introduction ..... 2**
- 1.1 Description ..... 2**
- 1.2 Constraints ..... 4**
  - 1.2.1 Implementation Constraints ..... 4
  - 1.2.2 Economic Constraints ..... 4
  - 1.2.3 Environmental Constraints ..... 4
  - 1.2.4 Social Constraints ..... 4
  - 1.2.5 Political Constraints ..... 5
  - 1.2.6 Ethical Constraints ..... 5
  - 1.2.7 Health and Safety Constraints ..... 5
  - 1.2.8 Sustainability Constraints ..... 5
- 1.3 Professional and Ethical Issues ..... 6**
- 2. Requirements ..... 6**
- 2.1 Non-Functional Requirements ..... 6**
  - 2.1.1 Usability ..... 6
  - 2.1.2 Reliability ..... 6
  - 2.1.3 Privacy and Security ..... 7
  - 2.1.4 Efficiency ..... 7
  - 2.1.5 Extensibility ..... 7
- 2.2 Functional Requirements ..... 7**
  - 2.2.1 System Functionality ..... 7
  - 2.2.2 User Functionality ..... 8
- 3. References ..... 9**

# 1. Introduction

Along with the developments in machine learning, most systems are transforming into automated models. Nevertheless, currently there is no available model to identify pollen types for palynology, the branch of biology which examines pollen.

According to our discussions with three palynologists from Zonguldak Bülent Ecevit and Ankara Universities, we have learned that palynologists in Turkey examine pollen samples manually and there is no such end product or any available ML model that automates this time-consuming procedure [1], [2], [3]. Furthermore, our discussions brought us to the realization that there is an immense need for a light microscope database of allergenic pollen nationwide [2], [3]. Finally, we have realized that there is a need for information and analysis sharing platform for palynologists as well as people with pollen allergies.

After long discussions with several palynologists, we have decided to create an ML model to classify allergenic pollen species in Turkey which can be expanded worldwide later, and furthermore create a database of these pollen for our model and future machine learning researches. Then, we will create a pollen map of Turkey where palynologists can exchange academic information. Moreover, people with pollen allergies can track the frequency, dates, and locations of these pollen. Finally, students and interns studying palynology have a hard time learning and distinguishing different types of pollen and hence their professor needs to spend considerable time helping students at the lab [1]. Therefore, our project will also help these students to study on their own by classifying their examples without the need of the professor.

## 1.1 Description

In our project, PolliVidis, there are three aspects that we will offer: classification and analysis of given pollen samples with a trained ML model, creation of a pollen database for further research, and construction of a pollen map of Turkey. After the project is finalized, any given pollen sample with allergenic pollen types will be classified autonomously. In addition, together with the pollen map, palynologists and academics will be able to exchange academic information. Moreover, people with pollen allergies will track the frequency dates, and locations of these pollen.

In the initial stage, there is a need for a large labeled pollen dataset which the ML model will be trained with. Unfortunately, all available datasets are designed for students and professors to learn the details of each pollen species. Thus, 2-5 images per pollen type are available [4]. We will constitute a light microscope pollen dataset for ML models

from scratch. After communication with Ankara University, Researcher Aydan Acar decided to help us constitute the dataset [3]. She has ready-to-be-photographed and purified prepareate of most common allergenic pollens in Turkey. We will take 200-300 photos of 25-35 pollen species to obtain 500-700 pollen photos per species. This photography procedure takes considerable time using a light microscope. Unfortunately, there is no available light microscope allergenic pollen dataset and we have to construct this dataset from scratch.

After the constitution of the dataset, we will design a neural network model to classify these allergenic pollens. The neural network will be a Convolutional Neural Network (CNN) type and it will be trained with our dataset. Several methods such as transfer learning and data augmentation will be used to increase the accuracy. After the training procedure, the model will be available through PolliVidis website for free pollen analysis for everyone.

Afterwards, the PolliVidis website will be built. The pollen map which will use Google Maps API will be added to the main page. We will design an interface so that everyone, without an account, will be able to access the pollen map and learn our analysis, allergenic pollen information based on time and location. Moreover, everyone can use our model to analyze their own pollen example via the website. However, only palynologists and academics can update the pollen map with their samples and uploading data to the map will require an academic account. Login will be required to protect the accuracy and the reliability of the map.

The analysis we plan to supply with our model will consist of pollen classification and counting. After uploading a sample, the user can learn the pollen types and their ratio (number) in their samples.

If academics and palynologists agree to share their sample, the pollen map will be updated with their analysis. Hence, everyone will be able to reach academic and statistical information using this platform and pollen map. Moreover, students will also be able to use the PolliVidis for educational purposes, such as uploading pollen photos and learning their type without consulting their instructor.

## 1.2 Constraints

### 1.2.1 Implementation Constraints

- PolliVidis will be developed as a web application.
- GitHub will be used for version control.
- GitHub Pages will be used to show project reports.
- PyTorch Python package will be used to construct the convolutional neural network.
- Google Maps API will be used to construct the pollen map.
- MySQL will be used to design the database.
- Python Django and React frameworks will be used for the website.
- Dataset to be trained will be constructed during the project.
- Dataset will be stored in the database of the project written in MySQL.
- Object Oriented Programming (OOP) paradigms will be followed.

### 1.2.2 Economic Constraints

- Website and analysis will be free for all users.
- Free libraries and APIs will be used.
- Free GitHub domain will be used for the website.
- Dijkstra Machine will be used for the server and database purposes.
- The future cost of domain, server, database, and maintenance of the website is not determined yet; it will fluctuate depending on the scale of the website.

### 1.2.3 Environmental Constraints

- Compressed images and efficient file formats should be used to emit 0.30 g of CO<sub>2</sub> to the atmosphere per page view, which is nearly 80% less than the average carbon emission of a website [5].

### 1.2.4 Social Constraints

- Sample and analysis upload to the pollen map will be allowed for academic staff only. Thus, anyone will not be able to upload random images to the map to prevent potential abuse.

- There should be a platform within the application where users can send their feedback that will be taken into consideration.
- There is no direct communication with other academics within the website. However, academic users can see each other's email addresses with their names.

### **1.2.5 Political Constraints**

- The website will not be related to any political activities and cannot be used for any political agenda.

### **1.2.6 Ethical Constraints**

- Personal information will not be shared via third parties.
- Data and analysis will be used to update the pollen map as long as the user allows it. Otherwise, any user can analyze their sample anonymously.
- While people with pollen allergies can only see the analysis results on the map, academics can also see the publisher of the analysis with their names and e-mail addresses.

### **1.2.7 Health and Safety Constraints**

- Correction and the accuracy of the pollen map will be protected by allowing only palynologists and academics to update the pollen map which will help pollen allergy patients to be informed about current allergenic pollen in Turkey.
- Every suggestion and analysis made by the PolliVidis should be considered as a suggestion for people with allergies.

### **1.2.8 Sustainability Constraints**

- Database and server maintenance will consume some energy which can be seen as the bare minimum impact for any website.
- Any type of printed or physical material will not be needed.
- Generated dataset will be published and eliminate the need of collecting the same samples again and again for future researches which saves human and electrical energy.

## **1.3 Professional and Ethical Issues**

- Personal information of a user should not be shared in any circumstances.
- Data and analysis will be used as long as the palynologist allows it.
- Anonymous analysis will be available.
- Personal information should be protected and stored as hashed in the database.
- Validity of each academic and palynologist should be checked.
- Reliability and correctness of shared data and analysis will be protected by allowing only palynologists and academics to update the map.
- All results and analysis in the website should be considered as suggestions and PolliVidis cannot be held responsible for any medical confusion.

## **2. Requirements**

### **2.1 Non-Functional Requirements**

#### **2.1.1 Usability**

- The application should be able to work on most search engines such as Safari, Chrome, Firefox, and Mozilla.
- The application should yield an analysis with the pollen samples taken under a light microscope within 2-3 seconds.
- Any user can access all features of PolliVidis in less than 3 clicks.
- Anonymous analysis without any login should be available.

#### **2.1.2 Reliability**

- The system should ensure that the pollen map data that it offers is reliable and obtained from a palynologist or academic's samples.
- The application should ensure reliable results (more than %90 accuracy) for the pollen classification.
- The application should not lose any pollen data unless the user deletes or doesn't let data to be added to the database.

### **2.1.3 Privacy and Security**

- The application should ensure that the user's data is safe by not storing their password directly but hashing it.
- The name of the palynologist, the location and the date of his/her samples will be shared with other academics as long as the palynologist allows it.
- Anonymous analysis without login is available and no personal information will be needed.

### **2.1.4 Efficiency**

- The webpage's loading time should not exceed 2 seconds which is the maximum loading time recommended by Google [6].
- Analysis will be conducted on the server rather than the user's own computer which should decrease memory usage.

### **2.1.5 Extensibility**

- The ML model can be improved in the future with more datasets available.
- The allergenic pollen in Turkey will be examined to scale the project. However, the pollen types can be increased in the future.
- Pollen map will be based on Turkey where the examined pollen is located. This map can be generalized to the World in the future.

## **2.2 Functional Requirements**

### **2.2.1 System Functionality**

- The application should detect and classify allergen pollen found in Turkey.
- The application should count the pollen and pollen density in the uploaded photo samples.
- The application should store the images that are uploaded by academics.
- The application should illustrate the pollen and its density on a map.
- The application should store users' names and passwords (as hashed).
- The application should allow new users to be created.
- The application should allow academics to look through the uploaded pollen database, publisher, and their contact info.
- The application should not permit irrelevant data to be uploaded.



- The application should be open to use without any registration; however, it should support academic accounts as well as admin accounts.

### **2.2.2 User Functionality**

- Academic users should be able to register and login to the system.
- Academic users should view other pollen images and analysis with publisher info.
- Any user should be able to view the pollen map, where the pollen density of Turkey is shown.
- Academic users should be able to upload pollen images whose data will be uploaded to the pollen map.
- Academic users should be able to specify whether the uploaded photo is appended to the pollen map or not.
- Any user should be able to see the count, density, and the species of the pollen from the analysis.

### 3. References

- [1] İ. Tekin, Ş. Alan, and T. Sarışahin, "Palynologists' Struggles," 15-Aug-2021.
- [2] İ. Tekin, Ö. Ünlüsoy, E. Ünal, U. A. Yürüten, E. G. Güliter, Ş. Alan, and İ. Ö. Tekin, "Pollen Projects," 23-Sep-2021.
- [3] İ. Tekin, Ö. Ünlüsoy, and A. Acar, "Ankara University and Pollen Project," 24-Sep-2021.
- [4] AutPal, "Palynological Database," *PalDat*. [Online]. Available: <https://www.paldat.org/>. [Accessed: 10-Oct-2021].
- [5] "How is your website impacting the planet?," *Website Carbon Calculator*, 02-Feb-2021. [Online]. Available: <https://www.websitecarbon.com/>. [Accessed: 08-Oct-2021].
- [6] "How fast should my website load?," *Blue Corona*, 20-Aug-2021. [Online]. Available: <https://www.bluecorona.com/blog/how-fast-should-website-be/>. [Accessed: 08-Oct-2021].