HACETTEPE UNIVERSITY
COMPUTER ENGINEERING DEPARTMENT

UNDERGRADUATE PROJECT PROGRESS REPORT

Project Name	Report Date
WEB PAGE COMPLEXITY STUDY	11.12.2020

Student Number(s)	Student Name(s)
21627543 21626901 21627802	Ece OMURTAY Deniz Ece AKTAŞ Ömer Bilal YAY
Supervisor(s)	**Company Representative(s)**
Murat AYDOS	

Project Coordinator	Report Approval
Date: _____	☐ Yes ☐ No If no, rational of rejection: _____

A. TECHNICAL PROGRESS

I. INTRODUCTION

Classification of the visual complexity level of web pages through CNN based learning methods.

The visual appearance of web pages affects the way a user will interact with the web page contents. The layout which shapes the visual characteristics of overall appearance is composed of various visual components such as texts, images, form elements, and white spaces. Moreover, the placement and visual presentation of these components influence the perceived visual complexity. At this point, the literature suggests that the visual complexity of a web page impacts the cognitive load and effort of the users when they interact with web pages.

This project investigates the use of convolution neural network-based deep learning methods for understanding the visual complexity level of the web pages. To do this, we firstly started gathering a base case by making a website and gathering user scores that provide an insight of how visually complex a website is from the screenshots we provided. Afterwards we will sector the base case and make a database and then we will find the feature vectors of websites and optimize the vectors and finally we are planning to work with different machine learning methods to classify the level of complexity of web pages based on dimensionally reduced vectors.

II. ARCHITECTURAL GOALS

First of all we had to implement a website that could work fort his we had to download asp.net packages, Visual Studio, SQL and Microsoft Access after software requirenments were filled we researched the topic and learnt some basics afterwards we will implemet machine learning but to do this part we need Ubuntu but we all have Windows so in Microsoft store there is an Ubuntu application that Works we will use that. As hardware weusing our own computers so we don't have very large memory spaces so we are using clouds fort his purpose.

III. ASSUMPTIONS AND DEPENDENCIES

Like we stated in previous section we are using our own computers so we are limited to the software systems that our computer works on and because our project is not supported by a company we don't have a budget so we can only use student packages of programs or free programs.

For spreading the survey we are also limited because of Covid we can't do the survey in school so we spread it virtually but some issues happen like not finding enough people etc.

IV. DECISIONS, CONSTRAINTS, AND JUSTIFICATIONS

To handle software issues, we are using free or student versions of programs and to handle hardware issues we are finding some solutions like ubuntu application or memory cloud systems.

V. ARCHITECTURAL MECHANISMS AND KEY ABSTRACTIONS

Architectural Mechanism 1 Having a working website to collect base-case for machine learning and dividing the database into labels; we implemented the website and it is working, we are collecting data but we have yet to divide the data into labels.

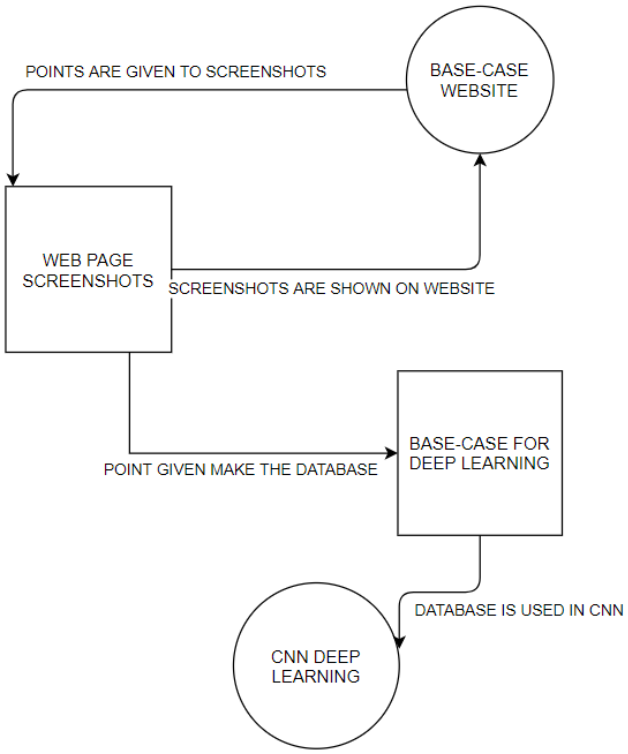
Architectural Mechanism 2 Finding the feature vectors of websites and optimize the vectors and finally we are planning to work with different machine learning methods to classify the level of complexity of web pages based on dimensionally reduced vectors

Keywords: CNN, Visual Complexity, Machine Learning, Asp.net,SQL

VI. LAYERS OF ARHITECTURAL FRAMEWORK

After we get all of our database we are going to label the database and divide it then start working with machine learning algorithms first and as the last step we will work with deep learning algorithms and decide on which algorithm gives the optimal results.

VII. ARCHITECTURAL VIEWS



B. PROJECT PROGRESS

I. CHANGES TO PROJECT PLAN

The change we made is; we had to delay machine learning part of the project for one month because the website implementation and data collecting parts took more time than anticipated.

II. PROGRESS OF PROJECT MILESTONES AND OBJECTIVES

Milestone #	Primary Objective	Due Date	Project Deliverable (if any)	Milestone Achieved?
1.	To finish the website for collecting data for base-case.	November 2020	A working website exists	Yes
2.	To get the database we will be using from the user inputs on the website and dividing them into labels.	December 2020	Database that will be used in machine learning and deep learning.	In process
3.	Project process evaluation and project process report delivery	December 2020	Project process reports will be submitted.	Yes
4.	To work on VGG and Resnet types of supervised CNNs with obtained labels in order to extract feature vectors	January 2020	Feature vector that is extracted from websites.	Not yet
5.	To apply a supervised technique named UMAP to reduce the dimensions in a supervised manner	January 2020	Optimized result is acquired.	Not yet
6.	Final project delivery and presentations	January 2020	Final project reports and presentations will be delivered.	Not yet

III. PROGRESS OF PROJECT PRACTICES AND MEASURES

Task #	Task Description	Responsibility	Start Date	Finish Date	Success Criteria	Task Succeeded?
1.	To get the base-case webpage development	Ece Omurtay Deniz Ece Aktaş Ömer Bilal Yay	October 2020	End of November 2020	Having a working website.	Yes
2.	Implementing the point acquired from website into database.	Ömer Bilal Yay	October 2020	December 2020	Having a database that is divided into labels.	In process
3.	Using the data base and implementing CNN and testing.	Ece Omurtay Deniz Ece Aktaş Ömer Bilal Yay	December 2020	January 2020	Getting the optimal results from the machine learning algorithms.	Not yet

Team Member	Task # Under Responsibility	Description of the Work Done

Ece Omurtay	1, 3	Did research on the topic. Worked with asp.net in Microsoft Visual Studio to implement a survey website for the purpose of getting base-case inputs for machine learning part. Spread the word about the survey and encouraged people to take the survey.
Deniz Ece Aktaş	1, 3	Did research on the topic. Worked with asp.net in Microsoft Visual Studio to implement a survey website for the purpose of getting base-case inputs for machine learning part. Spread the word about the survey and encouraged people to take the survey.
Ömer Bilal Yay	1, 2, 3	Did research on the topic. Implemented the database with SQL and afterwards helped with the implementation of the website by assisting with asp.net. Spread the word about the survey and encouraged people to take the survey.

IV. PROGRESS OF PROJECT BUDGET

We are going to use our own computers that we have and as for software we are planning to use Visual Studio, PyCharm and SQL but we will use the student packets of these software so we will not be doing any expenses and also because of the COVID-19 pandemic we will be working remotely so we will not be having any commuting expenses. As for income we don't have any income that is related to design project.

V. PROGRESS OF PROJECT RISKS

Risk Item #	Description	Probability	Effect	Did It Happen?	How did you (or will you) handle its occurrence? (Plan-B)
1.	If the user points are given in the start of the project is given with errors or given not seriously.	Possible	Results of study would be wrong	Yes	We implemented a confirmation system for points by taking five backup points and comparing them to other points.
2.	The machine learning methods we will try could give not the best results	Possible	Results wouldn't be optimal	No	We are planning to use more than one machine learning methods to see which one gives the best results.

VI. PROGRESS OF RESEARCH AND DEVELOPMENT (R&D) ACHIEVEMENTS

There are already works and researches done about visual complexities of websites but the difference of our study is that we will find the level of complexity of web pages specifically based on dimensionally reduced vectors using machine learning and deep learning is the innovative aspect of our project also we can add that rather than using a ready-made survey website we implemented our own to check the points so the inputs wouldn't have errors.

We made a website to get our base case from users for this purpose we used asp.net to develop our website and to gather the database we used SQL. Afterwards to implement methods like CNN we are planning to use python programming language so there haven't been any changes to technologies to develop.

VII. OVERALL PROGRESS OF YOUR PROJECT

Overall, we are learning a lot of stuff about computer vision, coding, website development, database management etc. especially with asp.net and SQL we had zero experience but after website development we understand this area much better but because of all of our other school work we take a little more time to archive our goals.