

Sketch2Code: From Sketch Design on Paper to Website Interface

Akash Wadje¹, Prof. Rohit Bagh²

^{1,2} *Computer Engineering, Sandip Foundation's SITRC, Nashik, India*

Abstract- Generally when any person want to build an website they have to hire an developer for that. For this user have to spend lot of money and time. But after paying so much money and time sometimes he doesn't get actual interface as they want. Currently, developer is the bridge between the user's sketches and the program code development .To overcome such type of problems and restriction Microsoft Artificial Intelligence Team introduces a platform named "SKETCH2CODE". Transforming a Handwritten Sketch (graphical user interface) created by a designer into computer code is a typical task conducted by a developer in order to build customized websites. In this paper, we show that different machine learning methods can be play major role to train a model end-to-end to generate code from a image given by user with over all different platforms support with help of browsers (i.e. iOS, Android and Windows).we present two approaches which automates this process, one using classical computer vision techniques, and another using a novel application of deep semantic segmentation networks. We uses a dataset of websites which is used to train and evaluate approaches.

Index terms- HTML, Artificial Intelligence, SKETCH2CODE, Graphical User Interface

I.INTRODUCTION

An early step in creating an application is to sketch a design on paper to know the structure of the interface. Designers face a challenge when converting their design into code; they have to spend their lot of time in designing of Graphical User Interface (GUI) rather than actual logics.

Majority of their time implementing the particular functionality and logic of the software they're building. Moreover, the computer languages designed to implement such GUIs are specific to each target runtime system; thus leading to tedious and repetitive work when the software being built is important to

run on multiple platforms using native technologies [1].

Earlier year, Microsoft introduces a platform called "Ink To Code" that translated drawings of a User Interface (UI) into code. It was very easy, user friendly: Even if you didn't have a lot of time or coding knowledge, you may still create a structure of a great idea.

Now Microsoft has taken the following step: "Sketch2Code", which allows users to upload a photo of a website or UI and have it translated into working code. The web-based program of Microsoft's Artificial Intelligence: For example, a computer-vision algorithm scans the photo and determines the placement of HTML elements, while a text-recognition module reads any text (for example, "I agree to terms and conditions" field).

Sketch2Code is more advanced version and user friendly than Ink to Code. You have to draw symbols to represent text fields and other specific elements. By Sketch2Code, you just needs to upload handwritten image on website with no special symbols necessary and the A.I. tries to take things from there [1].

We use an Computer Vision Technic to understand user requirement from paper to code generation. This way we can generate HTML code directly from a user's drawn image [1].

Problems in the domain of turning a design into code have been tackled before: SILK (Sketching Interfaces Like Krazy) turns digital drawings into application code using gestures; REMAUI (Reverse Engineering Mobile Application User Interfaces) converts high definition screenshots into mobile apps; PIX2CODE transforming user interface screenshots into computer code Many of these applications rely on classical computer vision techniques to perform detection [2].

The paper presented by the Tony Beltramelli in year 2017 regarding Generating Code from a Graphical

User Interface Screenshot a novel approach based on Convolutional and Recurrent Neural Networks allowing the generation of computer tokens from a single GUI screenshot as input, but the drawback of these system is that it is based on only software development platform not for web designing [2].

Also paper presented by Tuan Anh Nguyen, Christoph Csallner in 2015 regarding Reverse Engineering Mobile Application User Interfaces With REMAUI in which given input bitmap REMAUI identifies user interface elements such as texts, images, containers, and lists with the help of computer vision and optical character recognition (OCR) techniques. REMAUI does not integrate with tools that generate mobile application functionality either via Domain Specific Language (DSL) and it is also application based [3].

An paper presented by James A. Landay and Brad A. about research on Early stage of user interfaces they documented about the SILK technology in which designer roughly sketch the design of Interface that user want using electronic pad and stylus to quickly implement the User Interface. In these systems common uses of gestures in electronic sketch designing are facilitated. But it is not paper oriented sketch it is based on electronic sketch [4].

And finally paper presented by the Alexander Robinson in year 2018 describe Create an application which translates a sketched picture into a website, and to explore how deep learning compares to classical computer vision methods for this task which describe very effectively how to tackle existing system limitations using this new technology. We can say that deep learning methods for these task increases performance over classical computer vision techniques [2].

Since 3 year ago, Tony Beltramelli, co-founder and CEO of Uizard Technology, published a paper on how Generating Code From Graphical User Interface Screenshot he describe about the pix2code technology in which involves various machine learning disciplines and about computer vision capabilities[6].However, pix2code's great performance on its dataset it is not capable of generalizing real world examples.

As we discuss REMAUI is a technology which identifies user interfaces such as texts, images, containers using computer vision and optical character recognition (OCR) technologies. But,

REMAUI has challenges or problems when user wants to add new examples of classes which is time consuming also. Sometimes due to low level designs inclusion of style what user want not fully fulfilled [5].

As such, the goal of this system is to:

- a. Create an application which translates a sketch design directly into code;
- b. Compare classical computer vision techniques with deep learning methods in order to maximize performance.

Major Challenges involves are:

1. Detection of sketch perfectly to system avoiding all noises.
2. Layout adjustment with efficient use of training modules.
3. Identify elements and transform them into code.
4. Display the graphical user interface to user [2].

We understand the gap in research of applying deep learning techniques to low level design to code website.

We strongly believe that deep learning s suited to this type of problems and generate efficient output.

II. ARCHITECTURE PROCESSING

The architecture of Sketch2Code Consist of following elements:

- a. Microsoft Custom Vision Model: This model is trained with images of different handwritten design tagging information of most common HTML elements like Text Box, Label, Buttons and Images.
- b. Microsoft Computer Vision Service: To identify the text written into a design elements a Computer Vision Service is used.
- c. Azure Blob Storage: In these all the steps involved in HTML code generation process are stored including images, prediction result and layout grouping information.
- d. Azure Function: Act as a backend entry point that coordinate the generation process by interacting with its services.
- e. Azure Web App: The Front-end of portal to enable uploading a new design and see the generated HTML results [9].

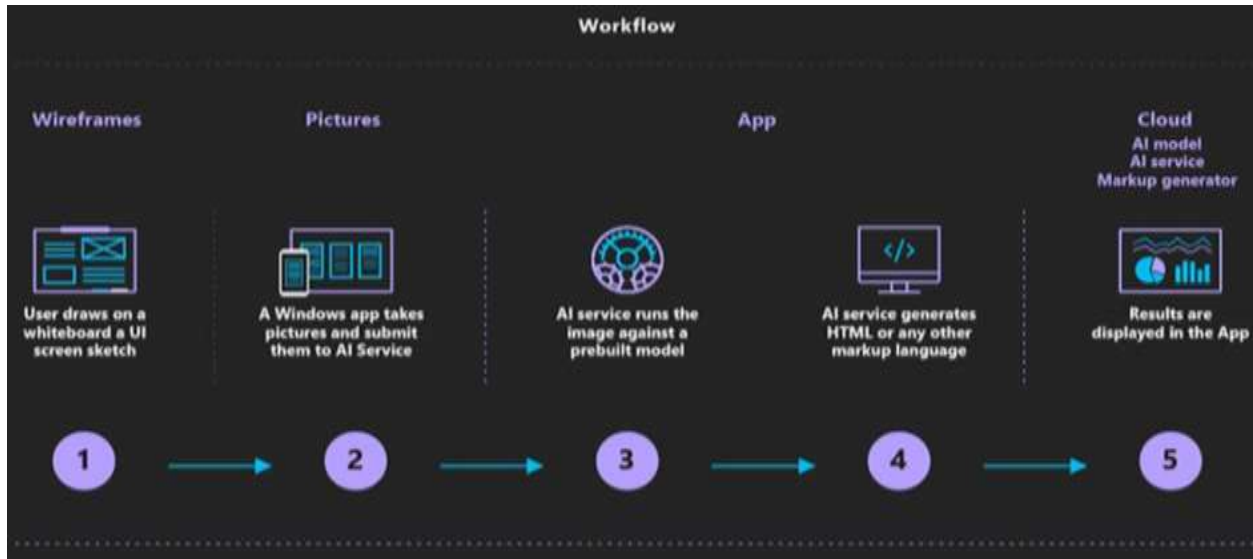


Figure 01: Procedure of Sketch2Code (From image upload to Azure Cloud Platform to Azure Cloud AI Web services)

A) MICROSOFT CUSTOM VISION MODEL

The training set used to create the sample model used in project is located in the Model folder. Each training image has unique identifier that matches information stored in dataset.json file. This file contain all the tags information used to train sample model. You can also create your Custom Vision Project at <https://customvision.ai>. After these you need to annotate a key and project Name to configure the Azure Function to call services [10].

B) COMPUTER VISION SERVICES

In this module we discuss about Computer Vision method which is used to converting images of sketches into HTML code. This approach involves Four Phases:

- Element Detection – Computer vision is used to detect and classify the size, position and type of all elements from sketch which is required for generation of HTML Code.
- Structure Detection – To build the hierarchical tree from list of all elements required for HTML code generation.
- Container Classification- classify types of container structure e.g. headers and footers.
- Layout Normalisation- Human errors during sketching Design are correct to produce required HTML code [2]

C) AZURE BLOB STORAGE

Azure Blob Storage allows to store an massive amount of unstructured data which does not need to be specific data model.

In our system we use it to store unstructured images format. The object stored in Blob Does not necessary have an extension.

A Blob contain many blocks but not more than 50,000 blocks per Blob.

The minimum size of block is 64KB and maximum is 100MB [11].

D) AZURE FUNCTION

The code for Azure Function is located in the sketch2Code API folders we can use this code to create our own function using configuration parameters :

- AzureWebJobsStorage :End point to Azure Storage
- ComputerVisionDelay : Time in millisecond to wait between call to Computer Vision Services.
- HandWrittenTextApiEndPoint : Endpoint to the Computer Vision Services.
- HandWrittenTextSubScriptionKey : key to access Computer Vision Service.
- ObjectDetectionItrationName : Name of training iteration to use in prediction model.
- Object DetectionPredictionKey: prediction key to access the Custom Vision Service.
- Object DetectionTraningKey: traning key for Custom Vision Service [10].

E) AZURE WEB APP :

In Sketch2Code.Web file contain code that used to implement front-end website. For this two parameter must be configured:

- Sketch2CodeAppFunctionEndPoint: Endpoint to Azure Function.
- StorageUrl: Url for the Azure function.

III. MACHINE LEARNING TECHNIQUES

ARTIFICIAL NEURAL NETWORK:

Artificial Neural Network (ANN) is basically computation model like human brains neurons. It is collection of input connection and output connection. A neuron computes weighted average of its input and pass through a non-linear function known as activation function.

MULTILAYER PERCEPTRON NETWORKS:

Multilayer Perception Network (MPN) it can distinguish data that is not linearly separable. MPN consist of three layers: Input layer , Hidden Layer , Output layer. When MPN is training input and output data pairs are set into MPN and algorithm updates weight between neurons to model the computation transformation which transform input data to output.

DATASET:

Deep learning methods requires a large datasets with number of samples. In dataset contain trained sketches and associated code. Human errors and different human opinions on how exactly to create an sketches could understand quality of dataset. We consider three options to create dataset:

- Finding Website and manually sketching them.
- Manually sketching website and building matching website.
- Finding website and automatically sketching them [2].

POST PROCESSING:

Note it, In this all system HTML code is not generated directly after just uploading image sketch it use intermediary Domain Specific Language(DSL), which is representation of structured sketches.

DSL is used because, representation of sketches not directly match to HTML we need translation step to convert the DSL into HTML.

IV. CONCLUSION

The goal of this thesis is to develop easy and efficient website from the sketch design of paper so that everyone can build website. In framework section explain how our framework has develop to use easily; by allowing images from web camera; by using common design symbols and training modules. Finally we design a technique to evaluate new system translate sketches into website.

REFERENCES

- [1] Nick Kolakowski – “Make codes from poor sketch” article on website creation.
- [2] Alexander Robinson : Generating a website from a paper Mockup - arXiv:1905.13750v1 [cs.CV] 9th May 2019 - University of Bristol MAY 2018
- [3] Tuan Anh Nguyen , Christoph Csallner: 2017 IEEE/ACM 39th IEEE International Conference on Software Engineering Companion 978-1-5386-1589-8
- [4] James A. Landay and Brad A. Myers – Carnegie Mellon University – Interactive Sketching for the early stages of User Interface Design
- [5] Mobile App Development – Tuan Anh Nguyen – Published by Lucy Marshall <https://slideplayer.com/slide/10896337/>
- [6] Thomas Claburn in San Francisco – “draw it and our AI do rest part” Article – 29th August 2018.
- [7] Tony Beltramelli - pix2code: Generating Code from a Graphical User Interface Screenshot: UIzard Technologies Copenhagen, Denmark - arXiv:1705.07962v2 [cs.LG] 19 Sep 2017
- [8] Microsoft official documentation on Sketch2Code on Microsoft AI.
- [9] Tara Jana Sr. Technical Product Marketing Manager, Artificial Intelligence - 27th August 2018
- [10] Documentation of Sketch2Code by AI_lab Microsoft on github.
- [11] Article of Nadeem Ahamed – “Azure blob storage vs file storage” – 23th August 2018