“FREEHAND DRAWN CIRCUIT RECOGNITION USING IMAGE PROCESSING”

Bachelor of Engineering
in
Computer Engineering

Submitted by

Ms. Shristi Mittal   16CE8010

Mr. Shubhamm Mohitte   13CE8013

Ms. Rhutuja Satpute   17CE5505

Guided by

(Dr. Leena Ragha)

D Y PATIL
RAMRAO ADIK
INSTITUTE OF
TECHNOLOGY
NAVI MUMBAI

Department of Computer Engineering

Ramrao Adik Institute Of Technology

Dr. D. Y. Patil Vidyanagar, Sector-7, Nerul, Navi Mumbai-400706.

(Affiliated to University of Mumbai)

April 2020

Freehand Drawn Circuit Recognition Using Image Processing

B.E. Project Report

Submitted in partial fulfillment of the requirements

For the degree of

Bachelor of Engineering
in
Computer Engineering

Submitted by

Ms. Shristi Mittal 16CE8010
Ms. Rhutuja Satpute 17CE5505
Mr. Shubhamm Mohitte 13CE8013

Guided by

(Dr. Leena Ragha)

Department of Computer Engineering

Ramrao Adik Institute Of Technology

Dr. D. Y. Patil Vidyanagar, Sector-7, Nerul, Navi Mumbai-400706.

(Affiliated to University of Mumbai)

April 2020

Ramrao Adik Institute of Technology

(Affiliated to the University of Mumbai)

Dr. D. Y. Patil Vidyanagar, Sector-7, Nerul, Navi Mumbai-400706.

CERTIFICATE

This is to certify that, the dissertation titled

“ Freehand Drawn Circuit Recognition Using Image Processing ”

is a bonafide work done by

Ms. Shristi Mittal 16CE8010
Ms. Rhutuja Satpute 17CE5505
Mr. Shubhamm Mohitte 13CE8013

and is submitted in the partial fulfillment of the requirement for the degree of

Bachelor of Engineering
in
Computer Engineering
to the
University of Mumbai

Supervisor	Co-Supervisor
(Dr. Leena Ragha)	**(Ms.Dhanashri Bhosale)**

Project Co-ordinator	Head of Department	Principal
(Mrs. Smita Bharne)	**(Dr. Leena Ragha)**	**(Dr. Mukesh D. Patil)**

Project Report Approval for B.E

This is to certify that the project entitled is a bonafide work done by **Shristi Mittal, Rhutuja Satpute and Shubhamm Mohitte** under the supervision of **Dr. Leena Ragha and co-supervision of Ms. Dhanashri Bhosale.** This dissertation has been approved for the award of **Bachelor's Degree in Computer Engineering, University of Mumbai.**

Examiners :

1. ....

2. ....

Supervisors :

1. ....

2. ....

Principal :

....

Date : .../.../......

Place : ....

Declaration

We declare that this written submission represents the ideas in our own words and where other's ideas or words have been included, we have adequately cited and referenced the original sources. We also declare that we have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in our submission. We understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Ms. Shristi Mittal	**16CE8010**	____________________
Ms.Rhutuja Satpute	**17CE5505**	____________________
Mr. Shubhamm Mohitte	**13CE8013**	____________________

Date : …/…/……

Abstract

Sketches are commonly used in the fields of engineering and architecture, especially for the early design phases. Engineers spend considerable time setting up initial designs using pencil and paper, and then redrawing them to any software. To solve this problem, the idea is to scan the circuit sketch drawn on the paper using a basic camera followed by detection and reconstruction of the circuit. The scanned image will be pre-processed and further segmented. The segmented image will be used to extract the features which are in turn given for classification. Recognizing sketches may seem so quick and intuitive to humans but it is really a big challenge for the machine. In this proposed work we aim to achieve a trainable electronic sketched circuit reconstructed with fast response time, high precision and simple extensibility to new components.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

In real life, it is often very handy to draw an electronic circuit with various components on paper. However, paper is not a reliable media for storing information. On the other hand, if we want to try things out and test if the sketched circuit is functional or not, it is impossible to realize on paper.In sketch reconstruction, a difficult task is to maintain a good balance between the ability to draw and the significance of recognition. A sketch recognition system would save a great deal of time for engineers to redraw these into professional applications. With the help of this system, they can provide scanned image as the input which is hand drawn on paper and get a reconstructed digital circuit. The project focuses on developing a system that will reduce human efforts to redraw the circuits on desktop and directly provide a scanned image as the input. Thus, by utilizing this, we can recognize the components and reconstruct the freehand image to a digital organized circuit. Digital circuits area unit less liable to noise or degradation in quality than analog circuits. It is conjointly easier using digitized circuits.
Electronics sector is immensely classified because it produces equipment for industries like telecommunications, electronic elements, producing and manufacturing equipment with growth within the semi-conductor business has become price over $ four hundred billion globally.[1] product created by this sector area unit wide utilized in a spread of client and industrial physics product, which incorporates PCBs (printed circuit boards) within those electronic product. PCB circuits can also cause errors like potential distance variations between the written lines, line thickness dissimilarities, false alarms owing to dirty atmosphere, etc. though the PCB business has enforced advances, these days constant reasonably errors continues to exist, and it's still developing examination determination. Hand drawn circuit recognition project may well be a

bridge between paper drawn circuit and generated digital circuit, which might eventually scale back time used for making a digital circuit for PCB.

1.2 Objectives

The main objectives of this paper is the recognition of hand-drawn circuit schematics with the application of digital image process. The person utilizing this project will discover hand drawn circuit elements and varied kinds of circuit diagrams into the machine.

1.3 Motivation

The fields of electronics have been relatively much slow in developing due to the in feasibility and inconvenience of its physical nature in the fast paced digital world. Engineers have to manually feed them to machine and convert them from an image diagram to machine understandable format. Our system aims to assist electrical engineers, students, examiners, professors in detailed detection and redrawing circuits of the free hand drawn electrical circuit making it easier to mobilize and practice them.

1.4 Organization Of Report

After the Abstract, the report is been organized into 8 chapters.
(i)The chapter 1 is Introduction, which contains sections like Overview of Project, Objectives and Motivation of the project.
(ii) The chapter 2 is Literature Survey, which contains sections like Survey of Existing System, its Limitations and the Problem Definition.
(iii) The chapter 3 named Proposal contains Proposed Work, Proposed Methodology and the Details of Hardware and Software Requirements.
(iv)The chapter 4, Planning and Formulation describes the schedule of the project with help of Gantt Chart.
(v) The chapter 5, Design of System, describes the overall design of the proposed system.
(vi) The chapter 6, gives the Results and their Discussion, explaining the implementation details, result analysis and the cost and benefit analysis.
(vii) The chapter 7 is Conclusion of the Project and t he Future Work based on it.

Chapter 2

Literature Survey

2.1 Survey of Existing System

The paper [1] summarizes the concept of machine recognition of hand-written circuit diagrams, having advantages of 92 percent accuracy for resistors and conductors. Small spots are often ignored in most systems of circuit recognition. However, in spite of the accuracy there is wrong classification of capacitors and voltage sources.

The paper [2] proposed a system of Hand-drawn digital logic circuit component recognition using Support Vector Machine and Fourier Descriptor. SVM simultaneously minimises the empirical classification error and maximises the geometric margin. The shape signature is selected which uses complex co-ordinates making it more accurate and efficient. An average of 83 percent circuit recognition accuracy is achieved. But the system has limited component detection.

Computer Vision Approach is extremely useful as it uses syntactic simulation in the paper[3]. It classifies components using a combination of invariant moments, scalar pixel-distribution features and vector relationships between straight lines in polygonal representations.

The paper [4] is an important project as it focuses on a new method using Artificial Neural Network (ANN) to make a machine that can directly read the electrical symbols from a hand drawn circuit image. It reduces the errors by combining with back propagation. Thus it modifies the weights of the various layers.

In [5] more components are detected and a combination of traditional pre-processing methods with modern novel methods are used to create a Netlist for simulation. We see that input symbols and data need not be provided.

The paper [6] gives us another way of recognizing freehand drawn circuits using Convolution Neural Networks (CNN). A network is trained using huge amounts of data set of 28 thousand images. It gives an accuracy of 91.85 percent in recognizing handwritten digits. but slow computing time makes it difficult.

Table 2.1 given below explains the analysis carried out among all the literature surveys. The application is developed in a way that it can analyze the components of the circuit and recognition whole circuit and perform simulation. They work by doing combination of feature extractions and segmentation and further classification using Support Vector Machine.

Table 2.1: Summary of Survey of Existing System

Sr.No	Name of Paper	Year	Methodology Used	Drawbacks
1.	Machine recognition of hand-written circuit diagrams	2000	Nearest Neighbour Classifier.	Misclassification of capacitors and voltage sources.
2.	Handwritten digital logic circuit recognition using SVM	2016	Support Vector Machine Fourier Descriptor	Input has only OR, NOT, AND gate components. Components need to be increased.
3.	Hand-drawn optical circuit recognition	2015	Artificial Neural Network through back propagation algorithm	It is not scalable in complexity.
4.	A novel method for circuit recognition through Image Processing.	2016	Traditional methods of pre-processing in OpenCV to get Netlist.	Lighting conditions and algorithm have not been specified.
5.	Handwritten digit recognition using OpenCV.	2019	Convolution Neural Networks.	Requires higher processing power and is slow due to the 5 hidden layers of CNN¿ Requires huge training dataset of 28k images.
6.	An electronic circuit diagram image recognizer.	1998	Feature point identification to guide a new line vectorization method.	Captures features with only very high dimensional points.

2.2 Limitations Of Existing Systems

From the literature and survey, we understood that there were some drawbacks and problems with the existing techniques:

1. Recognized circuit diagrams can be used as input to various simulator software.
2. Poorly chosen set of features which led to poor classification rates.
3. Wrong classification of capacitors and voltage sources.
4. Classification accuracy depends on the size of training set and drawing style of training images.
5. The classification methods used are not as efficient for scaling.

2.3 Analysis Of Literature Survey

On surveying the assorted systems, we've understood the downsides and understood the areas we'd like to work on to create an efficient and successful system. We understand that a robust data-set with varying features was missing in most systems leading to less accuracy of versatile input. We must aim to gather and build a versatile and real data from experts and amateurs as our project would be utilized by the users. Thus to form a decent size of training and testing dataset to figure on various classifiers. We see drawbacks in system because of ignorance of noise, small spots, shadows and ranging illumination, color and large size data where the photographs don't seem to be processed properly before giving as input causing poor feature extraction.[4] Pre-processing is an especially vital method to enhance the results of any classifier. The image must be clear, its parts and components precise for input. Therefore we have a tendency to learn that our focus should on pre-processing for the planned system. We realise that we want to settle on the proper set of features by creating a strong segmentation procedure.[2] We also conjointly realise that we want to detect all the vital components in an exceedingly circuit that may be helpful for a practitioner for simulation. The low accuracy within the existing systems is additionally because of feature extractor.[1] We see that HOG (Histogram of Oriented gradients) has worked more effectively as compared to SURF (Speeded up robust features) [2]. In the systems we tend to see that ANN, CNN and SVM dominate among others pretty much as good classifers. Neural networks like ANN and CNN perform higher on terribly massive datasets however they increase complexity. The system victimisation ANN gets too advanced and isn't scalable .[3] CNN provides the simplest accuracy of around 91[5] and SVM around 92 percent[1]. On additional survey we tend to realize that the system operating on CNN needs an enormous dataset and still performs slower. Our goal is to form this application for users in real life to be used existing hand held devices and not on GPU powered computers that CNN wouldn't be able to fulfil. SVM looks to be giving smart results with even tiny data-set. We understood that we must attempt SVM for the goal of our project. We additionally see SVM as a classifer which will be scaled. On the idea of those learnings and analysis we tend to create a proposed system.

Chapter 3

Project Proposal

3.1 Problem Statement

The existing detection methods are used to detect the documents comprise of text such as languages, mathematical symbols, digits, and medical symbols, while graphics detection has generated less interest than text detection. To start any project, it is necessary to draw its layout and then it is redrawn on software. It becomes very tedious to draw directly on any software and then judge its output and perform any alterations. So, this report focuses to minimize these efforts by developing a trainable system.

3.2 Scope

First of all it could be widely used at Institute level for branches like Electrical and Electronics for examiners to easily check the exam papers of those students.
Further it could generate a helping hand at Industry level for Designers of Electrical projects.
So our basic focus is on overall recognition of maximum components of the electrical circuits and their reconstruction.

3.3 Proposed Work

We propose a method to recognize hand drawn circuits using python programming language on the data set we gathered from various faculty and students of electronics and telecommunications department. The proposed work aims to recognize circuit components such as resistor, inductor, diode, capacitors, voltage source and ground.

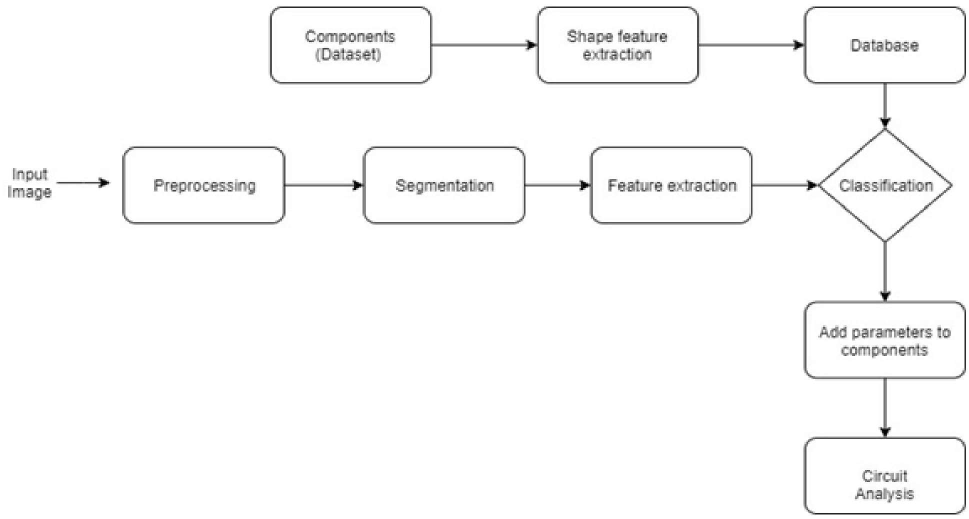


Figure 3.1: Proposed System Design

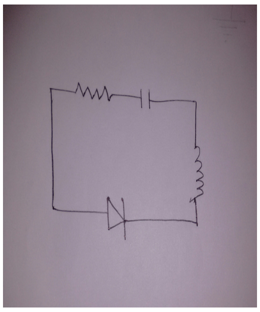
Fig 3.1 discusses the blocks or sub-parts of the projected drawing of a digital circuit of the hand-drawn circuit schematics system. It explains the detailed process of pre-processing followed by segmentation and feature extraction, also including classification, detection, and redrawing stages.

The scanned image is given as input to the pre-processing stage. The image ought to have a selected format like JPEG. Pre-processing is mostly accustomed to enhance the image quality, to shrink down the information size of the written circuit, and to make the data-set fit for input to further stages. In segmentation, the circuit is divided into parts and nodes by separating from connections with acceptable threshold acquired by the adaptive threshold technique. In the form feature extraction stage, each element is allotted a feature vector to spot it. This vector is employed to tell apart the element from another element. The detected circuit is then redrawn into a digital circuit.

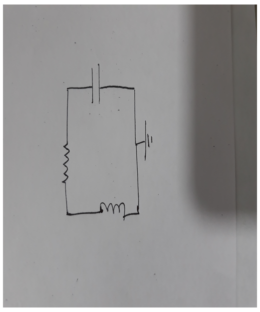
3.4 Proposed Methodology

3.4.1 Data-set

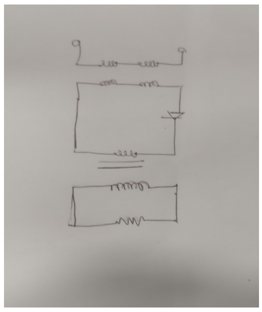
A decent quantity of data is needed to train Support Vector Machine. Most on-line data-sets don't seem to be accurate as they do not capture the varied features of hand-drawn circuits as real data does. We have collected the data from varied students and colleges from completely different departments in our university for flexibility and versatility in the data. The data is hand-drawn complete circuits. These pictures once captured square measure either blur, of low image quality, and additionally contain varied noise like paper lines, sunlight, and shadow. the photographs captured additionally could contain over one circuits and additionally text.



(a) Broken circuit



(b) Varying Illumination



(c) Blurred and multiple circuits

Figure 3.2: Versatility in data acquired

The training data-set consists of three files of 400 images for components for resistor, inductor, diode cropped from each circuit. These images are in jpeg format. The testing data-set has 100 images for the three components. The data-set also contains the components in all possible vertical and horizontal directions and the images are also rotated in three directions 90,180,270 to get all possible angles.

A circuit sketch is composed of component symbols and connection lines, each symbol needs a specified recognizer which is trained with a data set. Our interest is limited to "capacitor, inductor, diode, resistor, voltage source and ground components".

We also made some ideal conditions and assumptions listed below,

1. Connection lines are horizontal and vertical wires.
2. Open lines belong to a component symbol such as capacitor and ground.

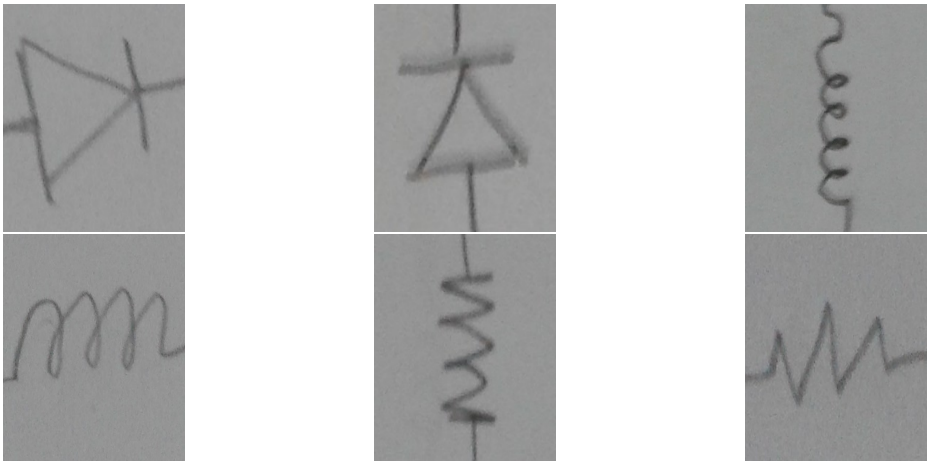


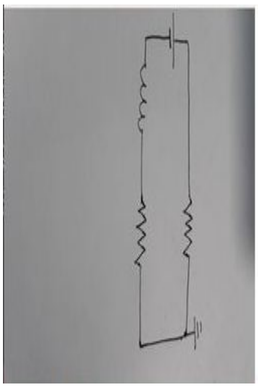
Figure 3.3: Images in training data-set

These assumptions make it easier to segment symbols and connection. Circuit recognition task consist of segmentation of component symbols, recognition of symbols and creating digital circuit.

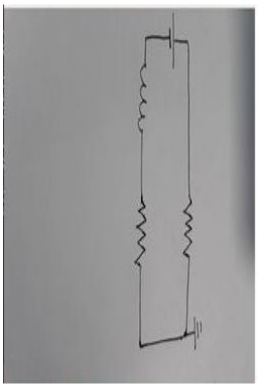
3.4.2 Pre-Processing

The proposed system takes the scanned image of hand-drawn circuits and supports the jpeg format. The images taken with a digital camera should be of a minimum of two pixels. The images of hand-drawn circuits obtained not only have lots of noise and illumination variations but are also of varied sizes and colors. Also, the circuits are almost scribbled thus having lots of errors, and hence they need to be pre-processed to create a clean and relevant data-set to feed as input to SVM for better classification rate. As our focus in this project is the outline and components of the circuit, the background data is irrelevant and hence increases size and noise when fed in for feature extraction and classification. Thus by applying image pre-processing techniques we have reduced the training time drastically and increased the accuracy of the classifier to a huge extent. Hence we have focused on pre-processing.

The pre-processing is composed of the following steps:
1. Resizing
The images that are captured using a camera are usually of different sizes due to the exposure and quality of the camera system. The pixel values are integers with values between 0 and 255. Most classifiers and neural networks use small weight values and inputs with large integer values can disrupt or slow down the process. First, these images are resized to a (100,100) pixel size which seemed the best small size to extract features from using HOG without too much loss and feed as input to SVM.



(a) Original image

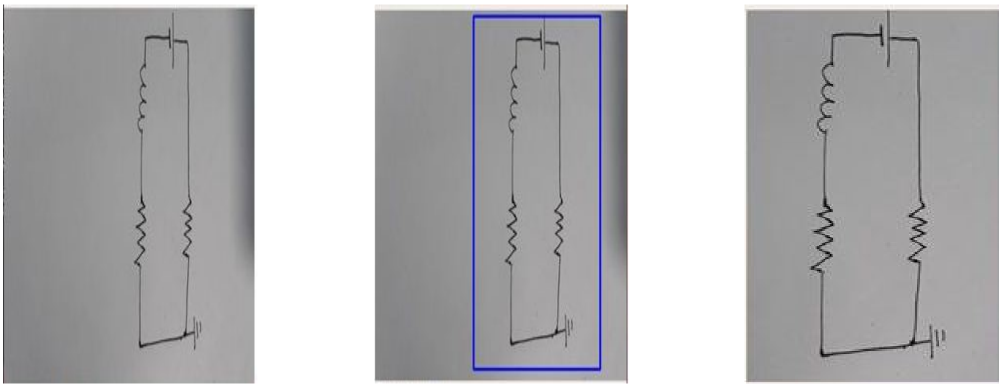


(b) Resized image

Figure 3.4: Resizing

2. Cropping

The image can then be cropped to avoid other text or circuits, as per the choice of the user. The user selects the start and endpoint using the cursor and selects the image he requires and crops it.

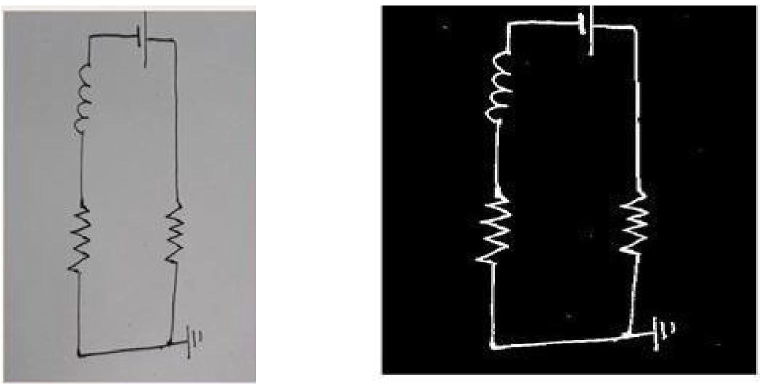


(a) Resized Image (b) To be cropped (c) Cropped image

Figure 3.5: Image cropping

3. Converting to Binary

We then convert the image from RGB to grayscale and further the image is converted into binary form to reduce the data for easier processing. As our images are versatile and have different lighting conditions in different areas, we use the Adaptive thresholding method as it gives us better results for images with varying illumination.

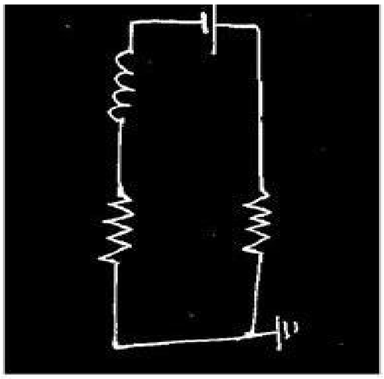


(a) Cropped image (b) Threshold Image

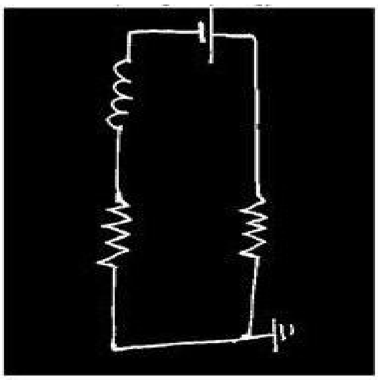
Figure 3.6: Image thresholding

4. Noise removal

Thresholding results in a binary image, but it consists of lots of noise that needs to be removed to not lead to confusion in ratio and cuts during feature extraction. Opening morphology was applied, it removes some of the foreground bright pixels from the edges and regions of important bright pixels. It preserves the circuit while removing the noise. We also tried erosion but it was destructive and had the disadvantage that if affect all regions with bright pixels (including the circuit) indiscriminately. The results are shown. The effect of the opening is easily visible and we see that all pixels which can be covered with structuring element within the bright region will be preserved but those that cannot be covered will be eroded.



(a) Threshold Image

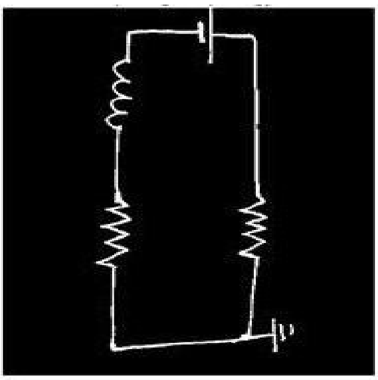


(b) Open morphology

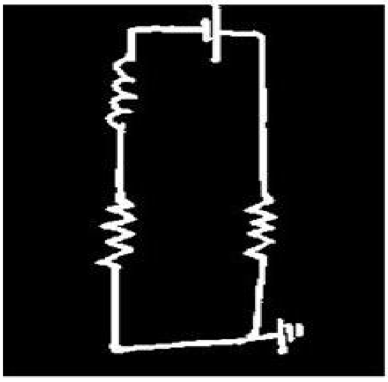
Figure 3.7: Noise removal

5. Dilation

Opening morphology thus left us with some broken circuit lines and thin joins that got eroded as noise. It becomes very important to join these lines to not confuse them with components and a ground to be confused as a diode. We use dilation to increase the circuit lines area so it joins the broken parts of the circuit and retaining the circuit while the noise is gone.



(a) Open morphology

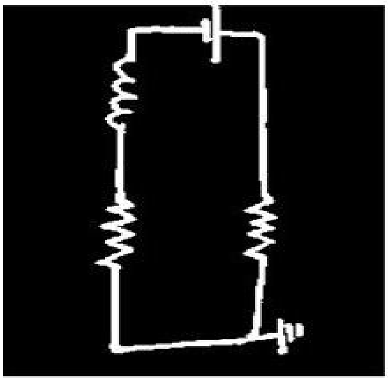


(b) Dilated

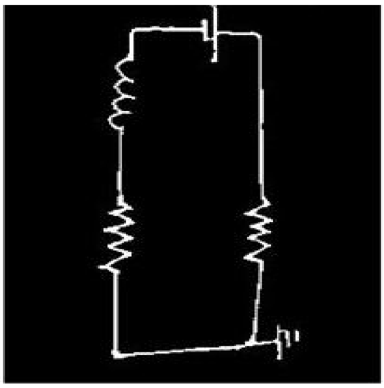
Figure 3.8: Dilation

6. Image thinning

The dilated image is too wide and may cause joining of two lines such as between capacitor which would cause confusion and reduce the efficiency of the classifier. The dilated image is then thinned, giving us a noiseless, connected circuit with thin lines.



(a) Dilated

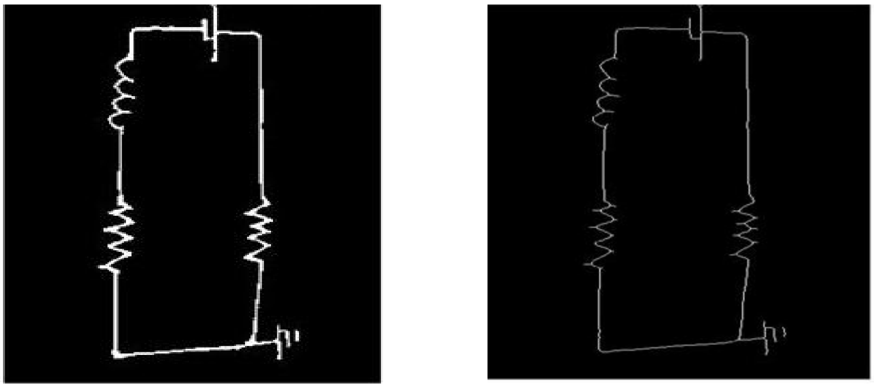


(b) Thinned

Figure 3.9: Thinning

7. Skeletonization

Ideally, the final image should have the minimum and most important data for better feature extraction. By using skeletonization, we convert the image into a simple yet compact representation of the circuit while capturing its essential features. This process converts the circuit lines in single-pixel width with high density, so we get a bright circuit and also reduce the size and unnecessary data away. By doing this we create better data and get the exact points where the circuit changes direction and also get the endpoints of the other components. Endpoints are extremely important as features during feature extraction for classification. Finally, with this step, we create the right dataset to give as input for segmentation.



(a) Thinned (b) Skeletonized

Figure 3.10: Skeletonization

3.4.3 Segmentation

Segmentation is the method of breaking apart the image into items that measure sufficiently small to be detected. Endpoints from pre-processing lead us to segment capacitor, voltage source, and ground symbols due to open lines they contain. We don't need a trained recognizer for these components because we know that the capacitor, voltage source, ground components differ from each other in the two cases, length ratio of lines, the number of lines they have.
We can identify them using these properties:
1. Capacitor has two lines with length ratio about 1
2. Voltage source has two lines with length ratio about 0.5.
3. Ground has different number of lines than other components.

The first assumption states that connection lines are horizontal and vertical lines. We detected these lines using the line segment detector algorithm [4]. Then, we removed detected connection lines and segmented components from adaptive thresholding done in the pre-processing stage. The morphological closing operation was applied to make the image suitable for the contour detection algorithm [5]. It returns separated foreground regions. We ignored small regions with thresholding contours using a threshold T by region area because they are probably remaining part of connection lines or noise due to illumination differences. The remaining regions are potential circuit components that can be identified by the recognition process.

3.4.4 Feature Extraction using HOG

After pre-processing and segmentation on the image of the circuit, features of every element are extracted. This step is the heart of the system as this step has a larger impact on the detection rate. This step helps to classify the elements according to their features. Feature extraction is that the name was given to a bunch of procedures for the measurement of the relevant form data contained during a pattern so the task of classifying the pattern is created simply by a proper procedure. Histogram Oriented Gradients[HOG] feature extraction techniques are generally suitable with the classifier Support Vector Machine(SVM).
Histogram Oriented Gradients[HOG] features are used to obtain feature vector of each training sample.

i) [HOG] algorithm uses gradient magnitude and direction of each pixel to create feature vector that describes a region of image.
ii) HOG divides the cells into 2*2 pixel cells.

iii) For a vector with x and y
Magnitude is calculated as

M = $\sqrt{dx^2_{(x,y)} + dy^2_{(x,y)}}$ (3.1)

M = $\sqrt{f^2_x + f^2_y}$ (3.2)

Direction is given by

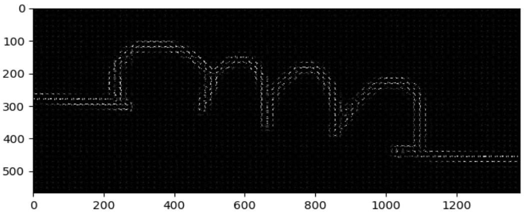
D = $\tan^-1 dy_{(x,y)}/dx_{(x,y}$(3.3)

D = $\tan^-1 f_y/f_x$(3.4)

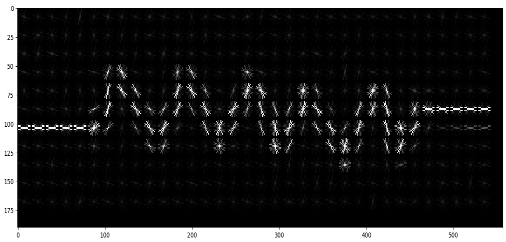
where x and y are components of the vector;
M = magnitude, D = direction; dx and dy = differentiation factors of x and y;
fx = function of x;fy = function of y;

It is a rotation and scale invariant feature description algorithm. Histogram Oriented Gradients[HOG] feature vector and label of each training sample was used to train Support Vector Machine(SVM).

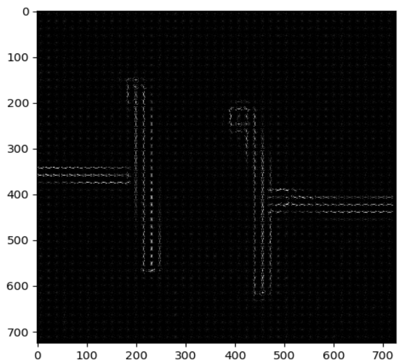
The features extracted by HOG for the various components are given below.
We study that the features for resistor are continuous curved lines and for resistor differ as they are sharped hill continuous lines. Similarly the magnitude and direction for capacitor is two lines with ratio of 1.0 and for voltage source a ratio of about 0.5 units. The features extracted for ground are three small lines and the diode can be classified having a closed triangular shape feature. As these images may have have cuts, noise and the lines may be slant, we have trained the extractor by giving it a number of circuits for it to better understand and identify the lengths of the ratio.



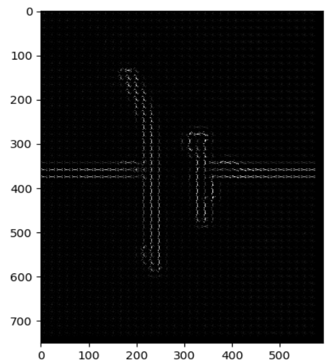
(a) HOG of the inductor



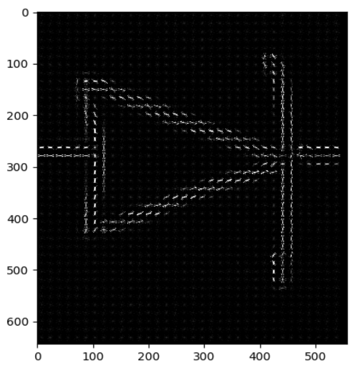
(b) HOG of the resistor



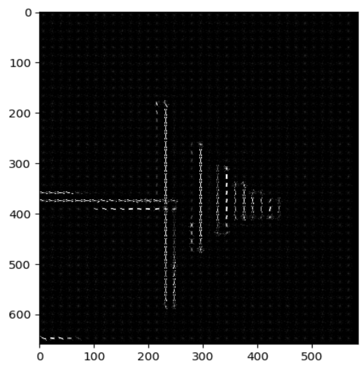
(c) HOG of the capacitor



(d) HOG of the voltage



(e) HOG of the diode



(f) HOG of the ground

3.4.5 Classification using Support Vector Machine

Classification is the process of labelling of objects into one of a number of predefined categories. We have a supervised data. Thus we can classify using supervised learning algorithms. As analysed in the literature survey CNN also seem to be competitive with SVM. But we see the benefits of SVM and see it could solve our problem. So we try this classification model. Support Vector Machine (SVM) minimizes error in the empirical classification and maximizes the geometrical margin.
The original SVM algorithm was invented by Vladimir N. Vapnik and Alexey Ya. Chervonenkis in 1963. In machine learning, support-vector machines (SVMs, that only rely on support-vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. A support-vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks like outliers detection. Experimental results show that SVMs achieve significantly higher search accuracy than traditional query refinement schemes after just three to four rounds of relevance feedback. This is also true for image segmentation systems, including those using a modified version SVM that uses the privileged approach. Also it benefits from proper feature selection and extraction techniques.

We use linear Support Vector Machine(SVM), to get an output of three different classes - Resistor, Inductor and Diode. The algorithm we used for SVM is as follows:
1. Import the training dataset.
2. After calculating histogram of features, scale the features.
3. Fit the SVM classifier to the training set.
4. Save the SVM

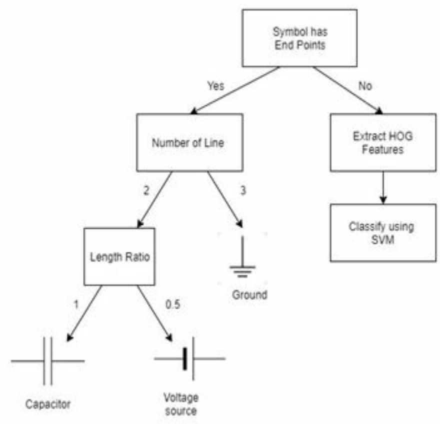


Fig:2.5 : Classification Flowchart

The Fig 2.5 represents the flow of classification of our system. Each component is checked if it has end points. If yes, the number of lines are checked in the component and it is further classified as ground for 3 lines. For 2 lines it is classified as capacitor or voltage source on the basis of length ratio.If the component has no end points it is given the HOG to extract features and further to classify using SVM.

3.4.6 Detection and Reconstruction

The detection goal is to interpret the circuit taken from the scanner. In the identification of the same sequence of operations and redraw based on the Support Vector Machine(SVM) classifier labels.

3.5 Details of Hardware/Software Requirement

1. Hardware requirements :

Processor	Intel i3/ Intel i5
RAM	4 GB
Scanner	Camera any mega-pixel length

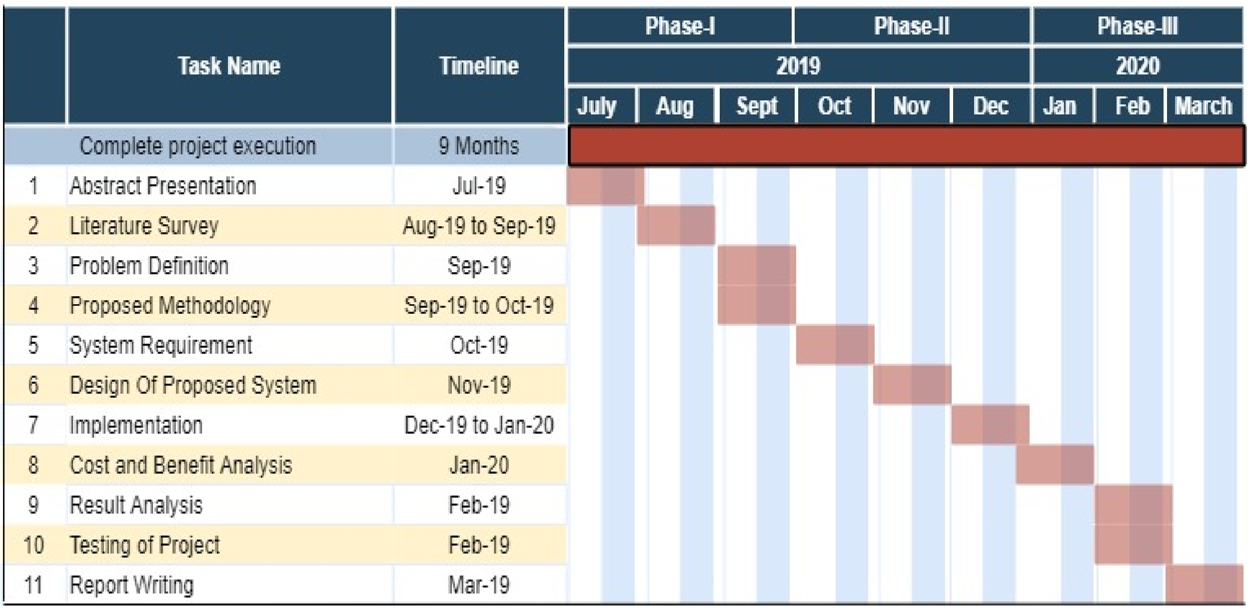
2. Software requirements :

Operating System	Ubuntu
Interpreter	Python
Python Libraries with their versions	imutils=0.5.2 numpy=1.14.0 opencv-python=3.4.0.12 pylsd=0.0.2 scipy=0.18.1 scikit-learn=0.14 scikit-image=0.14

Chapter 4

Planning And Formulation

4.1 Schedule for Project



	Task Name	Timeline
	Complete project execution	9 Months
1	Abstract Presentation	Jul-19
2	Literature Survey	Aug-19 to Sep-19
3	Problem Definition	Sep-19
4	Proposed Methodology	Sep-19 to Oct-19
5	System Requirement	Oct-19
6	Design Of Proposed System	Nov-19
7	Implementation	Dec-19 to Jan-20
8	Cost and Benefit Analysis	Jan-20
9	Result Analysis	Feb-19
10	Testing of Project	Feb-19
11	Report Writing	Mar-19

Figure 4.1: Schedule For Project

Fig 4.1 explains to us the Gantt Chart for the project. It shows the schedule and the flow of work of the project. The above timeline explains the start from our study on the Abstract till the Testing and Report Writing of the project.

4.2 Detail Plan of Execution

Following is the detailed plan of execution

1. Defining the problem.
2. Studying the previous programs and papers and surveying them.
3. Understanding the need and usability in real life.
4. Figuring out the objectives and scope of the roposed system.
5. Developing Block Diagram.
6. Collecting dataset from faculties and students from various branches such as electronics,electrical and computer science.
7. Capturing all the images and cropping them to make individual input files for Resistor, Inductor, Diode.
8. Developing Flowchart for entire process.
9. Writing, Compiling actual program, testing and debugging.

Chapter 5

Design of System

5.1 Design Diagram with Explanation

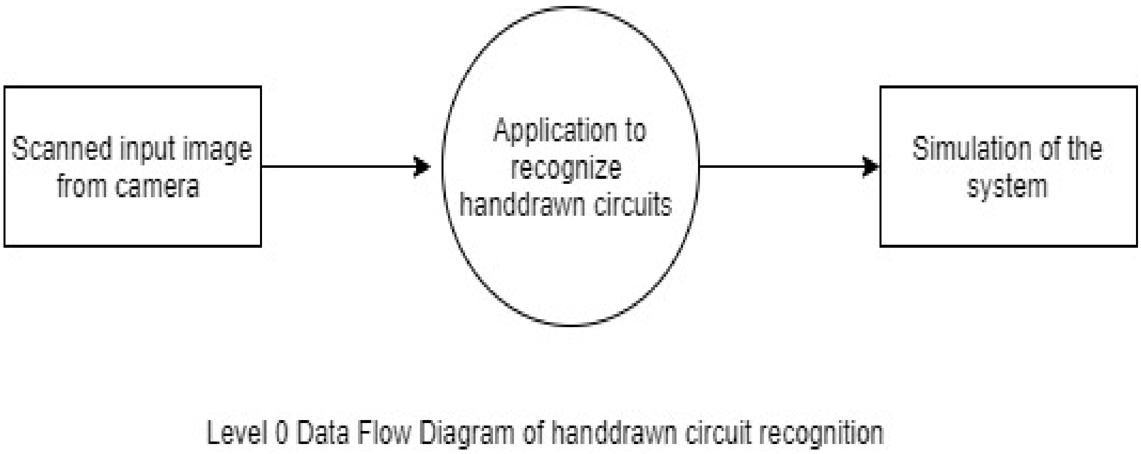


Figure 5.1: DFD level 0

Fig 5.1 explains the Data Flow Diagram (level 0) of the system we designed. The scanned image from the camera is sent to the application to recognize the hand-drawn circuits and can be further sent to simulator.

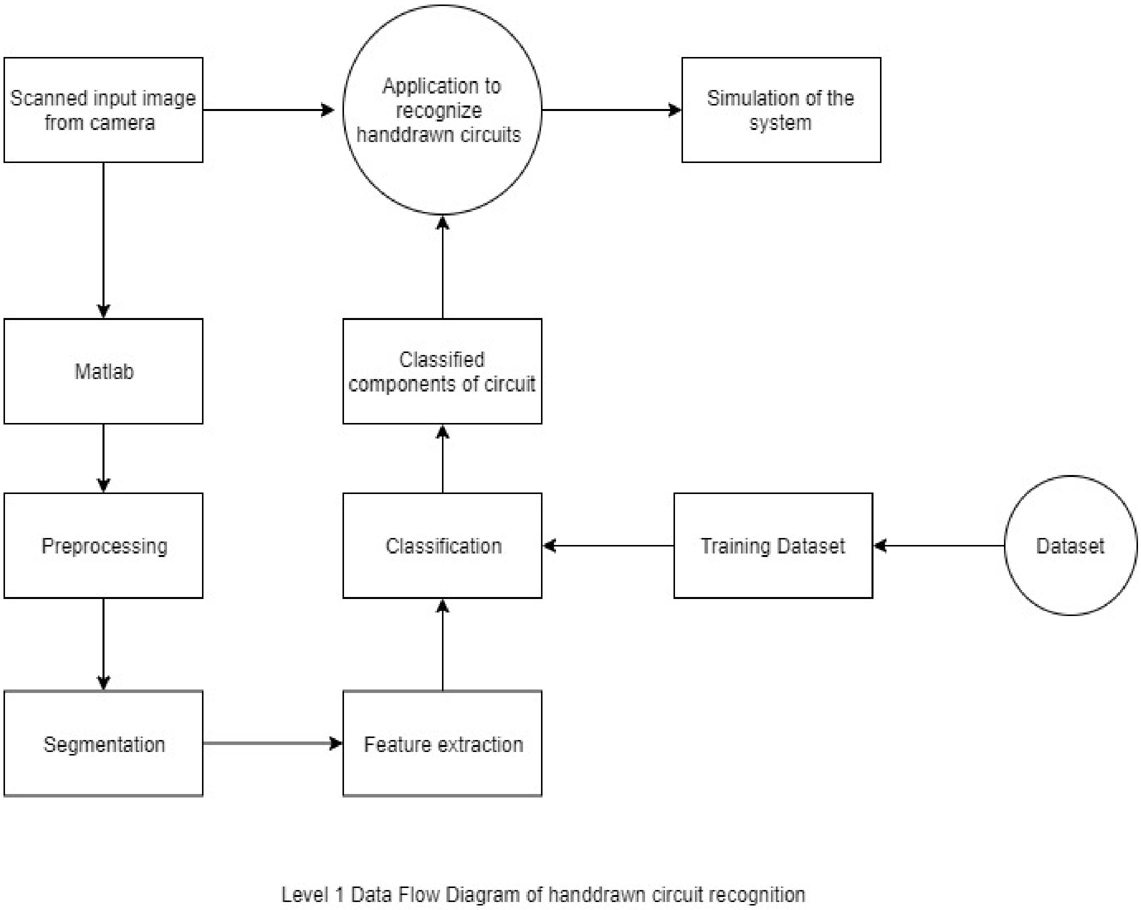


Figure 5.2: DFD level 1

Fig 5.2 is the next step of the Data Flow Diagram (level 1) which explains what happens inside the system. We would have the training dataset which consists of all images classified. Scanned image is further preprocessed and segmented. Next step is to extract features which are further compared with the dataset available and based on that classification takes place. Once the classification of components takes place, digital circuit is drawn.

Chapter 6

Results And Discussion

6.1 Implementation Details

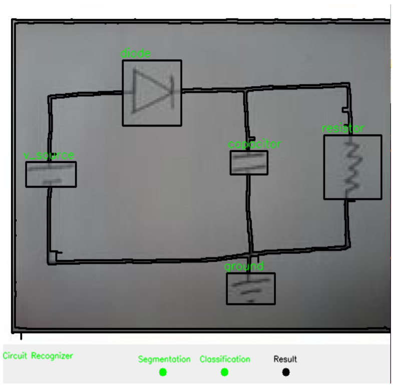
On applying the proposed methods and implementing them, we go through the following process. We see that jpeg and png images are scanned using the camera and given as input to the image processor. Usually images have noise or often are captured blurry, so to rectify these cases, we pre-process the image, i.e, we turn image into gray-scale, convert into binary using adaptive threshold.
Further in this binary image if the circuit points are not joined or any there's noise in the binary image, they are removed using morphological operations such as Opening, Closing, Erosion and Dilation. After that we convert it into a single pixel wide length binary circuit image also called as skeletonized image, which is helpful for segmentation of the components further.
Once the components are segmented from the circuit, we match it with the features extracted by using HOG (histogram of oriented gradients) while training the Support vector machine classifier.The SVM classifier classifies the image and detects the right component. After that the user gets the output and the circuit is reconstructed.

6.2 Result Analysis

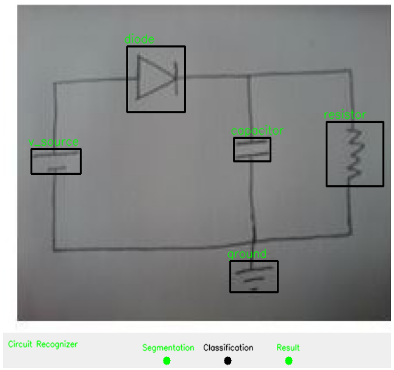
6.2.1 Overview

The system working on parameters depending on features would provide the result indicating the separation of components and wired nodes from the circuits. All the processes from image acquisition to classification of circuit and redrawing a digital circuit will be in a same place.

The resultant data will be the product of duly trained Support Vector Machine(SVM) which is trained under several similar data to provide precise accurate results more than 90% . Hence, a system is developed to detect the components, nodes and gates of the provided drawn circuit. Some of our results :



(a) Classification of Components



(b) Final Result of reconstructing lines of the circuit

Figure 6.1: Result

6.2.2 Confusion between components

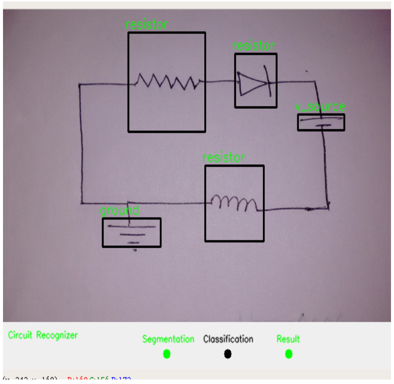
We observe that a few components are confused by our SVM trainer and results in wrong detection.

1. The resistor gets confused as inductor due to the unclear round curves of the inductor and the lines of the resistor and vice-versa.

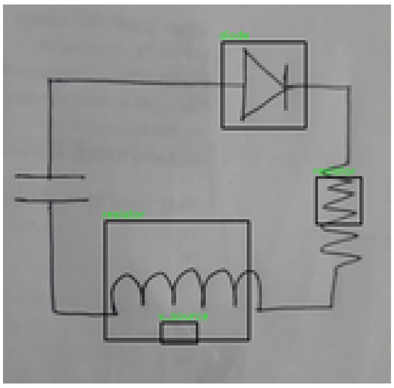
2. Similarly the resistor and ground are confused in some cases due to the three continuous lines and the spaces not being accurately understood and hence recognising them as the resistor's up and down lines.

3. The diode is a vertical line with a triangle, this triangle gets confused as the resistor's hill peak line.

Hence, these confusions get reduced and the results keep getting accurate as our SVM trainer gets more experience.



(a) Diode and Inductor detected as Resistor



(b) Inductor detected as Resistor and Voltage Source

Figure 6.2: Result showing confusion between components

6.2.3 Performance Measure

Table 6.1: Performance measure of component detection

	Resistor	Inductor	Diode	Capacitor	Voltage Source	Ground	Not Dectected
Resistor	90	5	0	0	0	5	10
Inductor	10	90	0	0	0	0	10
Diode	5	0	92	0	0	3	8
Capacitor	0	0	0	96	4	0	4
Voltage Source	0	0	0	5	95	0	5
Ground	5	0	0	0	0	95	5

The table for the percentage of performance of the recognised components as we have used training and testing ratio as 80 is to 20, i.e, 80- training data and 20-testing data which gave 95% accuracy for Resistors 90% for Inductors and 100% for Diode,90% for capacitor, 95%for voltage source and 95% for ground as well.

6.3 Cost Analysis

The Cost and Benefit is calculated for an organisation to understand how much they would have to spend on the project and the benefit of the project, to see if the benefits outweigh the costs.

We have computed the cost estimation based on three parameters
1. Cost estimation of Lines of Code using COCOMO model
2. Cost of Hardware
3. Cost of Software
1. COCOMO cost estimation

COCOMO is a constructive cost model. COCOMO uses a basic regression formula with parameters that are derived from historical project data and current as well as future project characteristics. It estimates the most important cost i.e the the efforts cost and schedule for software projects.

We can calculate COCOMO cost with

$$EffortsappliedE = a * (KLOC)^b \tag{6.1}$$

$$DevelopmentTimeD = c * (E)^d \tag{6.2}$$

$$PeopleRequiredP = E/D \tag{6.3}$$

where a=2.4, b=1.05, c=2.5, d=0.38 respectively LOC=1000 approximately KLOC of projects =1.0

$$Effort = 2.4 * (1.0)^(1.05) = 2.4 \tag{6.4}$$

$$NominalDevelopmentTime = 2.5 * (2.4)^(0.38) = 3.5months \tag{6.5}$$

We get the Development Time =3.5 months But we also need to understand the efforts in studying and researching the literature and creating the methodology which takes the same amount of time as coding.

Therefore Time For the Project =7 months

2. Cost of Hardware

The hardware required is a computer of RAM and processing power which costs approximately INR 35,000/- The digital camera required could be the normal camera in the computer or mobile device or basic digital camera.

3. Cost of Software
The software used are the frameworks, libraries and the programming language python which are all available for open source use for free.

6.4 Benefit Analysis

1.This website is useful for students to get to know about their errors they may have made while drawing the circuit.
2.This can be also used by teachers while they are examining the answwr sheets.
3. It will reduce human efforts to redraw the circuits on desktop and directly provide a digital reconstructed image.

Chapter 7

Conclusion And Future Work

7.1 Conclusion

Hand-drawn circuits are used widely in electrical and electronics engineering. The digital recognition of hand-drawn circuit can save a lot of time from using tiring and complex software. The proposed method is believed to be significant providing an efficient system for reconstructing hand drawn electrical circuits into the system improving the accuracy. Many new components can be added to the database and this method can be made very effective for a large number of electrical components.

7.2 Future Work

1. This project can be enhanced by adding simulation of the circuits feature, which will give the results of the input values hand written by the user.
2. The GUI can be incorporated into an Android Application for easy use on mobile devices.
3. This project can also be enlarged to take more components and work on logic gates.

Chapter 8

References

[1] B. Edwards and V. Chandran, “Machine Recognition of Hand-Drawn Circuit Diagrams”, Research Concentration in Speech, Audio and Video Technology, School of Electrical and Electronic Systems Engineering Queensland University of Technology, Brisbane, Australia.

[2] Mayuri D. Patare (Post graduate student) and Madhuri S. Joshi (Professor of CSE Department), “Hand-drawn Digital Logic Circuit Component Recognition using SVM”, Jawaharlal Nehru Engineering college, Aurangabad, Maharashtra, India.

[3] Williams, Kyle; Filho, Milton Ribeiro; and Renshaw, Megan, ”Automatic contextual recognition of hand-drawn content”, Technical Disclosure Commons, March, 2018.

[4] Srikanth Sridar and Krishnan Subramanian, Circuit Recognition Using Netlist Proceedings of the 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013).

[5] Momina Moetesum1, Syed Waqar Younus1, Muhammad Ali Warsi1 and Imran Sid-diqi1. Segmentation and Recognition of Electronic Components in Hand-Drawn Circuit Diagrams. Research Article: EAI.EU, 10.4108/eai.13-4-2018.154478.

[6] Yasser Alginahi (2010), ”Preprocessing Techniques in Character Recognition”, Character Recognition, Minoru Mori (Ed.), ISBN: 978-953-307-105-3, Talibah University, Kingdom of Saudi Arabia.

Appendices

Appendix A

Plagiarism Report

S26_Freehand_Drawn_Circuit_Recognition_Using_Image_Proce...

ORIGINALITY REPORT

15%	8%	7%	12%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Ramrao Adik Institute of Technology Student Paper	4%
2	www.aiktcdspace.org:8080 Internet Source	2%
3	en.wikipedia.org Internet Source	1%
4	Sachin Bende, Rajashree Shedge. "Dealing with Small Files Problem in Hadoop Distributed File System", Procedia Computer Science, 2016 Publication	1%
5	Submitted to Edith Cowan University Student Paper	1%
6	Rabbani, Mahdi, Reza Khoshkangini, H.S. Nagendraswamy, and Mauro Conti. "Hand Drawn Optical Circuit Recognition", Procedia Computer Science, 2016. Publication	1%
7	par.cse.nsysu.edu.tw Internet Source	1%

Appendix B

Paper Publication

Shristi Mittal, Rhutuja Satpute, Shubhamm Mohitte, Ms. Dhanashri Bhosale, ”Freehand Drawn Circuit Recognition and Reconstruction”, submitted/ presented at International Conference on Automation, Computing and Communication 2020 (ICACC-2020). The Venue of ICACC 2020 is Ramrao Adik Institute of Technology, Sector-7, Pad. Dr. D.Y. Patil Vidyapeeth, Nerul, Navi Mumbai- 400706.

Freehand to Digital Circuit Reconstruction Using HOG and SVM

Ms. Shristi Mittal
Department of Computer Engineering
Ramrao Adik Institute Of Technology
(Affiliated to University of Mumbai)
Dr. D. Y. Patil, Vidyanagar,
Sector 7, Nerul, Navi Mumbai,
Maharashtra

Ms. Rhutuja Satpute
Department of Computer Engineering
Ramrao Adik Institute Of Technology
(Affiliated to University of Mumbai)
Dr. D. Y. Patil, Vidyanagar,
Sector 7, Nerul, Navi Mumbai,
Maharashtra

Mr. Shubhamm Mohitte
Department of Computer Engineering
Ramrao Adik Institute Of Technology
(Affiliated to University of Mumbai)
Dr. D. Y. Patil, Vidyanagar,
Sector 7, Nerul, Navi Mumbai,
Maharashtra

Dr. Leena Ragha
Department of Computer Engineering
Ramrao Adik Institute Of Technology
(Affiliated to University of Mumbai)
Dr. D. Y. Patil, Vidyanagar,
Sector 7, Nerul, Navi Mumbai,
Maharashtra

Ms. Dhanashri Bhosale
Department of Computer Engineering
Ramrao Adik Institute Of Technology
(Affiliated to University of Mumbai)
Dr. D. Y. Patil, Vidyanagar,
Sector 7, Nerul, Navi Mumbai,
Maharashtra

***Abstract*—Sketches are commonly used in the fields of engineering and architecture, especially for the early design phases. Engineers spend considerable time setting up initial designs using pencil and paper, and then redrawing them to any software. To solve this problem, the idea is to scan the circuit sketch with android device which is drawn on the paper and translate it into standard layouts and run circuit simulations. The scanned image will be pre-processed and further segmented. The segmented image will be used to extract the features which are in turn given for classification. Recognizing sketches may seem so quick and intuitive to humans but it is really a big challenge for the machine. In this proposed work we aim to achieve a trainable electronic sketched circuit recognizer with fast response time, high precision and simple extensibility to new components.**

I. Introduction

A. Overview:

In real life, it is often very handy to draw an electronic circuit with various components on paper. However, paper is not a reliable media for storing information. On the other hand, if we want to try things out and test if the sketched circuit is functional or not, it is impossible to realize on paper.In sketch recognition, a difficult task is to maintain a good balance between the ability to draw and the significance of recognition. A sketch recognition system would save a great deal of time for engineers to redraw these into professional applications. With the help of this system, they can provide scanned image as the input which is hand drawn on paper and directly get the results.

The project focuses on developing a system that will reduce human efforts to redraw the circuits on desktop and directly provide a scanned image as the input. Thus, by using this we can redraw to check whether the system is functioning properly or not and also to checkout the results of the circuit drawn.Digital circuits area unit less liable to noise or degradation in quality than analog circuits. It is conjointly easier to perform error detection and correction with digital signals using digitized circuits.

Electronics sector is immensely classified because it produces equipment for industries like telecommunications, electronic elements, producing and manufacturing equipment with growth within the semi-conductor business has become price over $ four hundred billion globally.[1] product created by this sector area unit wide utilized in a spread of client and industrial physics product, which incorporates PCBs (printed circuit boards) within those electronic product. PCB circuits can also cause errors like potential distance variations between the written lines, line thickness dissimilarities, false alarms owing to dirty atmosphere, etc. though the PCB business has enforced advances, these days constant reasonably errors continues to exist, and it's still developing examination determination. Hand drawn circuit recognition project may well be a bridge between paper drawn circuit and generated digital circuit, which might eventually scale back time used for making a digital circuit for PCB.

Similarly, this project may well be helpful for education systems like for physics department for teaching students, process those fashioned circuits yet as any drawn circuit may well be digitized to see if the operating of the circuit is correct.

B. Objectives:

The main objectives of this paper is that the recognition of hand-drawn circuit schematics with the application of digital image process. The person utilizing this project will discover hand drawn circuit elements and varied kinds of circuit diagrams into the machine.

C. Motivation:

The fields of electronics have been relatively much slow in developing due to the in feasibility and inconvenience of its physical nature in the fast paced digital world. Engineers have to manually feed them to machine and convert them from an image diagram to machine understandable format. Our system aims to assist electrical engineers, students, examiners, professors in detailed detection and redrawing circuits of the free hand drawn electrical circuit making it easier to mobilize and practice them.

D. Survey of Existing systems:

The paper of B.Edwards and V.Chandran with concept of ”Machine recognition of hand-written circuit diagrams”[1] has advantage of 92 percent accuracy for resistors and conductors. Small spots are often ignored in most systems of circuit recognition. However, in spite of the accuracy there is wrong classification of capacitors and voltage sources.

”Proposed a system of Hand-drawn digital logic circuit component recognition using Support Vector Machine and Fourier Descriptor” by Mayuri D. Patare and Madhuri S. Joshi has Support Vector Machine(SVM)[2] which at the same time minimizes error in the empirical classification and maximizes the geometric margin. The shape signature is chosen which makes it more precise and effective using complex coordinates. The overall accuracy of circuit detection is reached by 83 per cent. But the device does have minimal detection of components.

”Computer Vision Approach is extremely useful as it uses syntactic simulation in ”Automatic contextual recognition of hand-drawn content” by Williams, Kyle, Filho, Milton Ribeiro, Renshaw and Megan [3]. It classifies parts employing a combination of invariant moments, scalar pixel-distribution options and vector relationships between straight lines in two-dimensional figure representations.

”Circuit Recognition using netlist proceedings” by Srikanth Sridar and Krishnan Subramanian [4] is a vital project because it focuses on a brand new methodology victimization Artificial Neural Network (ANN) to form a machine that may directly browse the electrical symbols from a hand drawn circuit image. It reduces the errors by combining with back propagation. so it modifies the weights of the assorted layers.

”Segmentation and Recognition of Electronic components in hand-drawn circuit diagrams” by Momina Moetesum, Syed Waqar Younus, Muhammad Ali Warsi and Imran Sid-diqi [5] concluded that more components are detected and a combination of traditional pre-processing methods with modern novel methods are used to create a Netlist for simulation. We see that input symbols and data need not be provided.

”Pre-processing Techniques in Character Recognition” by Yasser Alginahi[6] gives us another way of recognizing freehand drawn circuits using Convolution Neural Networks (CNN). A network is trained using huge amounts of data set of 28 thousand images. It gives an accuracy of 91.85 percent in recognizing handwritten digits. but slow computing time makes it difficult.

E. Analysis of existing systems:

On surveying the assorted systems, we’ve understood the downsides and understood the areas we’d like to work on to create an efficient and successful system. We understand that a robust data-set with varying features was missing in most systems leading to less accuracy of versatile input. We must aim to gather and build a versatile and real data from experts and amateurs as our project would be utilized by the users. Thus to form a decent size of training and testing dataset to figure on various classifiers. We see drawbacks in system because of ignorance of noise, small spots, shadows and ranging illumination, color and large size data where the photographs don’t seem to be processed properly before giving as input causing poor feature extraction.[4] Pre-processing is an especially vital method to enhance the results of any classifier. The image must be clear, its parts and components precise for input. We realise that we want to settle on the proper set of features by creating a strong segmentation procedure.[2]

The low accuracy within the existing systems is additionally because of feature extractor.[1] We see that HOG (Histogram of Oriented gradients) has worked more effectively as compared to SURF (Speeded up robust features) [2]. In the systems we tend to see that ANN, CNN and SVM dominate among others pretty much as good classifers. Neural networks like ANN and CNN perform higher on terribly massive datasets however they increase complexity. The system victimisation ANN gets too advanced and isn’t scalable .[3] CNN provides the simplest accuracy of around 91[5] and SVM around 92 percent[1]. On additional survey we tend to realize that the system operating on CNN needs an enormous dataset and still performs slower. Our goal is to form this application for users in real life to be used existing hand held devices and not on GPU powered computers that CNN wouldn’t be able to fulfil.

II. Working

A. Dataset Used :

A decent quantity of data is needed to train Support Vector Machine. Most on-line data-sets don’t seem to be accurate as they do not capture the varied features of hand-drawn circuits as real data does. We have collected the data from varied students and colleges from completely different departments in our university for flexibility and versatility in the data. The data is hand-drawn complete circuits. These pictures once captured square measure either blur, of low image quality, and additionally contain varied noise like paper lines, sunlight, and shadow. the photographs captured additionally could contain over one circuits and additionally text.

The training data-set consists of three files of 400 images for components for resistor, inductor, diode cropped from each circuit. These images are in jpeg format. The testing data-set has 100 images for the three components. The data-set also contains the components in all possible vertical and horizontal directions and the images are also rotated in three directions 90,180,270 to get all possible angles.

A circuit sketch is composed of component symbols and connection lines, each symbol needs a specified recognizer which is trained with a data set. Our interest is limited to "capacitor, inductor, diode, resistor, voltage source and ground components". We also made some ideal conditions and assumptions listed below,
1. Connection lines are horizontal and vertical wires.
2. Open lines belong to a component symbol such as capacitor and ground.

These assumptions make it easier to segment symbols and connection. Circuit recognition task consist of segmentation of component symbols, recognition of symbols and creating digital circuit.

B. Pre-processing :

The proposed system takes the scanned image of hand-drawn circuits and supports the jpeg format. The images taken with a digital camera should be of a minimum of two pixels. The images of hand-drawn circuits obtained not only have lots of noise and illumination variations but are also of varied sizes and colors. Also, the circuits are almost scribbled thus having lots of errors, and hence they need to be pre-processed to create a clean and relevant data-set to feed as input to SVM for better classification rate. As our focus in this project is the outline and components of the circuit, the background data is irrelevant and hence increases size and noise when fed in for feature extraction and classification. Thus by applying image pre-processing techniques we have reduced the training time drastically and increased the accuracy of the classifier to a huge extent. Hence we have focused on pre-processing.

The pre-processing is composed of the following steps:

1. Resizing
The images that are captured using a camera are usually of different sizes due to the exposure and quality of the camera system. The pixel values are integers with values between 0 and 255. Most classifiers and neural networks use small weight values and inputs with large integer values can disrupt or slow down the process. First, these images are resized to a (100,100) pixel size which seemed the best small size to extract features from using HOG without too much loss and feed as input to SVM.

2. Cropping
The image can then be cropped to avoid other text or circuits, as per the choice of the user. The user selects the start and endpoint using the cursor and selects the image he requires and crops it.

3. Converting to Binary
We then convert the image from RGB to grayscale and further the image is converted into binary form to reduce the data for easier processing. As our images are versatile and have different lighting conditions in different areas, we use the Adaptive thresholding method as it gives us better results for images with varying illumination.

4. Noise removal
Thresholding results in a binary image, but it consists of lots of noise that needs to be removed to not lead to confusion in ratio and cuts during feature extraction. Opening morphology was applied, it removes some of the foreground bright pixels from the edges and regions of important bright pixels. It preserves the circuit while removing the noise. We also tried erosion but it was destructive and had the disadvantage that if affect all regions with bright pixels (including the circuit) indiscriminately. The results are shown. The effect of the opening is easily visible and we see that all pixels which can be covered with structuring element within the bright region will be preserved but those that cannot be covered will be eroded.

5. Dilation
Opening morphology thus left us with some broken circuit lines and thin joins that got eroded as noise. It becomes very important to join these lines to not confuse them with components and a ground to be confused as a diode. We use dilation to increase the circuit lines area so it joins the broken parts of the circuit and retaining the circuit while the noise is gone.

6. Image thinning
The dilated image is too wide and may cause joining of two lines such as between capacitor which would cause confusion and reduce the efficiency of the classifier. The dilated image is then thinned, giving us a noiseless, connected circuit with thin lines.

7. Skeletonization
Ideally, the final image should have the minimum and

most important data for better feature extraction. By using skeletonization, we convert the image into a simple yet compact representation of the circuit while capturing its essential features. This process converts the circuit lines in single-pixel width with high density, so we get a bright circuit and also reduce the size and unnecessary data away. By doing this we create better data and get the exact points where the circuit changes direction and also get the endpoints of the other components. Endpoints are extremely important as features during feature extraction for classification. Finally, with this step, we create the right dataset to give as input for segmentation.

C. Segmentation :

Segmentation is the method of breaking apart the image into items that measure sufficiently small to be detected. Endpoints from pre-processing lead us to segment capacitor, voltage source, and ground symbols due to open lines they contain. We don't need a trained recognizer for these components because we know that the capacitor, voltage source, ground components differ from each other in the two cases, length ratio of lines, the number of lines they have.
We can identify them using these properties:
1. Capacitor has two lines with length ratio about 1
2. Voltage source has two lines with length ratio about 0.5.
3. Ground has different number of lines than other components.

The first assumption states that connection lines are horizontal and vertical lines. We detected these lines using the line segment detector algorithm [4]. Then, we removed detected connection lines and segmented components from adaptive thresholding done in the pre-processing stage. The morphological closing operation was applied to make the image suitable for the contour detection algorithm [5]. It returns separated foreground regions. We ignored small regions with thresholding contours using a threshold T by region area because they are probably remaining part of connection lines or noise due to illumination differences. The remaining regions are potential circuit components that can be identified by the recognition process.

D. Feature Extraction :

After pre-processing and segmentation on the image of the circuit, features of every element are extracted. This step is the heart of the system as this step has a larger impact on the detection rate. This step helps to classify the elements according to their features. Feature extraction is that the name was given to a bunch of procedures for the measurement of the relevant form data contained during a pattern so the task of classifying the pattern is created simply by a proper procedure. Histogram Oriented Gradients[HOG] feature extraction techniques are generally suitable with the classifier Support Vector Machine(SVM). Histogram Oriented Gradients[HOG] features are used to obtain feature vector of each training sample.
i) [HOG] algorithm uses gradient magnitude and direction of each pixel to create feature vector that describes a region of image.

ii) HOG divides the cells into 2*2 pixel cells.

iii) For a vector with x and y
Magnitude is calculated as

$$M = \sqrt{dx^2_{(x,y)} + dy^2_{(x,y)}} \quad (1)$$

$$M = \sqrt{f_x^2 + f_y^2} \quad (2)$$

Direction is given by

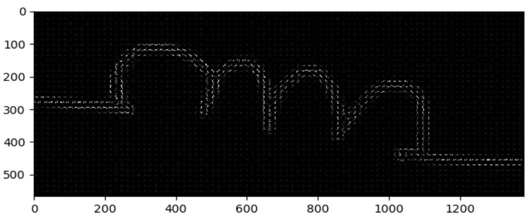
$$D = \tan^{-1} 1 dy_{(x,y)}/dx_{(x,y} (3)$$

$$D = \tan^{-1} 1 f_y/f_x (4)$$

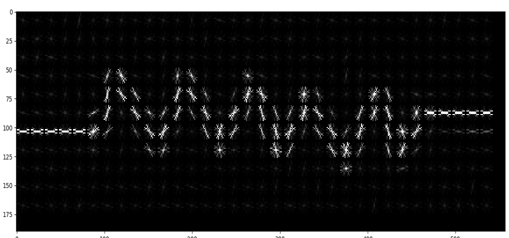
where x and y are components of the vector; M = magnitude; D = direction; dx and dy = differentiation factors of x and y;
fx = function of x;fy = function of y;

It is a rotation and scale invariant feature description algorithm. Histogram Oriented Gradients[HOG] feature vector and label of each training sample was used to train Support Vector Machine(SVM).

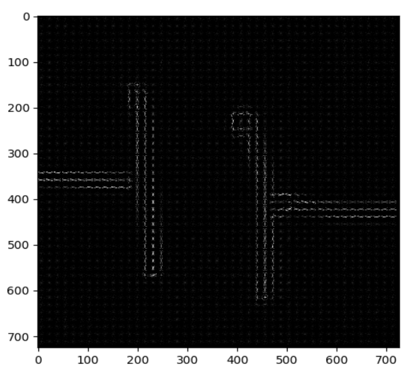
The features extracted by HOG for the various components are given below.
We study that the features for resistor are continuous curved lines and for resistor differ as they are sharped hill continuous lines. Similarly the magnitude and direction for capacitor is two lines with ratio of 1.0 and for voltage source a ratio of about 0.5 units. The features extracted for ground are three small lines and the diode can be classified having a closed triangular shape feature. As these images may have have cuts, noise and the lines may be slant, we have trained the extractor by giving it a number of circuits for it to better understand and identify the lengths of the ratio.



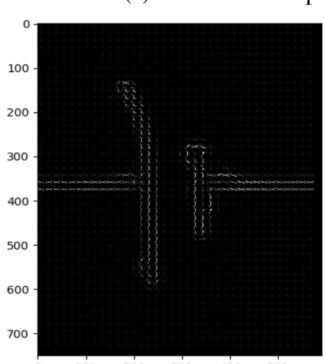
(a) HOG of the inductor



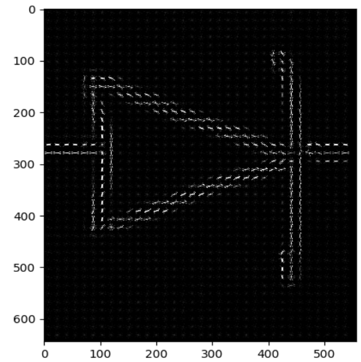
(b) HOG of the resistor



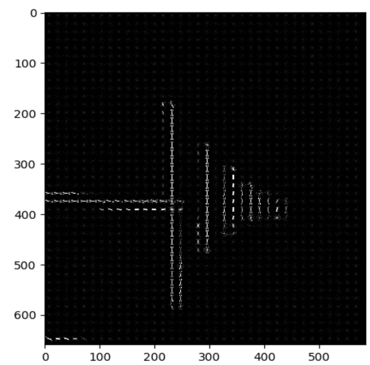
(c) HOG of the capacitor



(d) HOG of the voltage



(a) HOG of the diode



(b) HOG of the ground

E. Classification :

Classification is the process of labelling of objects into one of a number of predefined categories. We have a supervised data. Thus we can classify using supervised learning algorithms. As analysed in the literature survey CNN also seem to be competitive with SVM. But we see the benefits of SVM and see it could solve our problem. So we try this classification model. Support Vector Machine (SVM) minimizes error in the empirical classification and maximizes the geometrical margin.
The original SVM algorithm was invented by Vladimir N. Vapnik and Alexey Ya. Chervonenkis in 1963. In machine learning, support-vector machines (SVMs, that only rely on support-vector networks) are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. A support-vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks like outliers detection. Experimental results show that SVMs achieve significantly higher search accuracy than traditional query refinement schemes after just three to four rounds of relevance feedback. This is also true for image segmentation systems, including those using a modified version SVM that uses the privileged approach. Also it benefits from proper feature selection and extraction techniques.

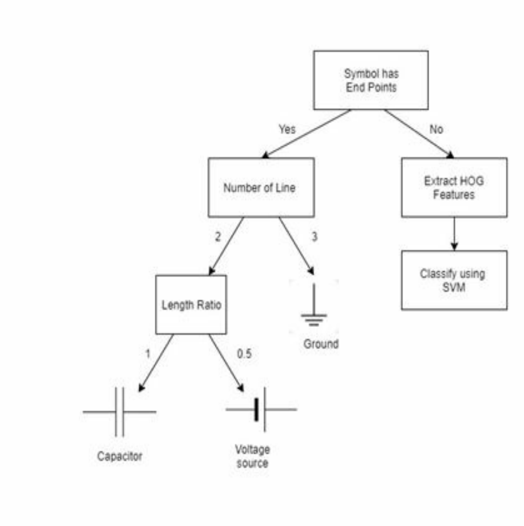


Fig. 3: Classification Flowchart

We use linear Support Vector Machine(SVM), to get an output of three different classes - Resistor, Inductor and Diode. The algorithm we used for SVM is as follows:
1. Import the training dataset.
2. After calculating histogram of features, scale the features.
3. Fit the SVM classifier to the training set.
4. Save the SVM The Fig 3 represents the flow of classification of our system. Each component is checked if it has end points. If yes, the number of lines are checked in the component and it is further classified as ground for 3 lines. For 2 lines it is classified as capacitor or voltage source on the basis of length ratio.If the component has no end points it is given the HOG to extract features and further to classify using SVM.

F. Detection and Redrawing :

The detection goal is to interpret the circuit taken from the scanner. In the identification of the same sequence of operations the circuit is detected by comparing the extracted features with simulation of the trained classifier. Then, the components are identified and redrawn based on the Support Vector Machine(SVM) classifier.

III. Results and Discussion

A. Overview:

The system working on parameters depending on features would provide the result indicating the separation of components and wired nodes from the circuits. All the processes from image acquisition to classification of circuit and redrawing a digital circuit will be in a same place. The resultant data will be the product of duly trained Support Vector Machine(SVM) which is trained under several similar data to provide precise accurate results more than 90% . Hence, a system is developed to detect the components, nodes and gates of the provided drawn circuit.

Some of our results :

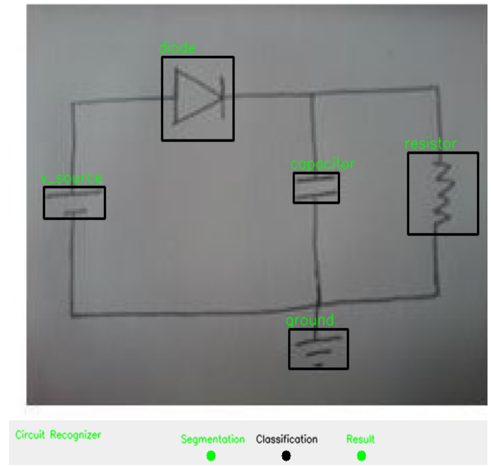


Fig. 4: Classification of components

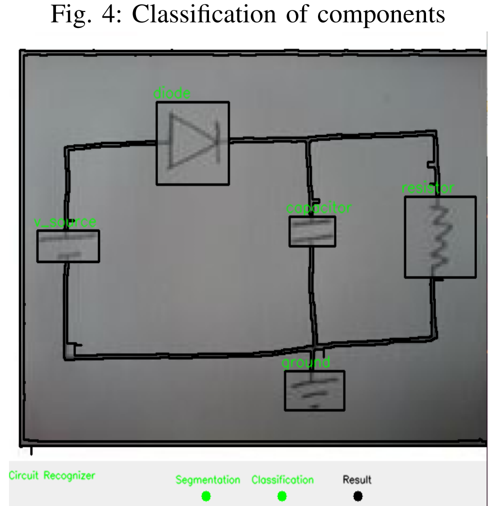


Fig. 5: Final Result of reconstructing lines of the circuit

B. Confusion between components :

We observe that a few components are confused by our SVM trainer and results in wrong detection.

1. The resistor gets confused as inductor due to the unclear round curves of the inductor and the lines of the resistor and vice-versa.

2. Similarly the resistor and ground are confused in some cases due to the three continuous lines and the spaces not being accurately understood and hence recognising them as the resistor's up and down lines.

3. The diode is a vertical line with a triangle, this triangle gets confused as the resistor's hill peak line.
Hence, these confusions get reduced and the results keep getting accurate as our SVM trainer gets more experience.

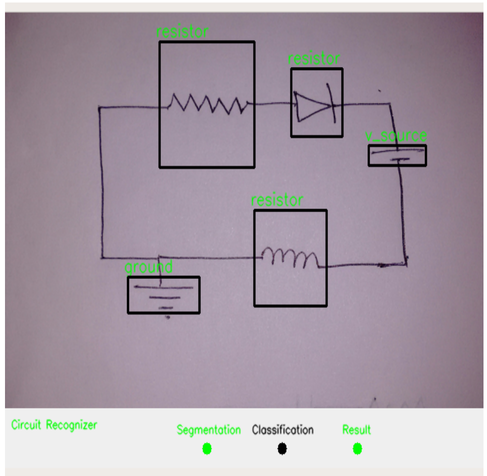


Fig. 6: Diode and Inductor detected as Resistor

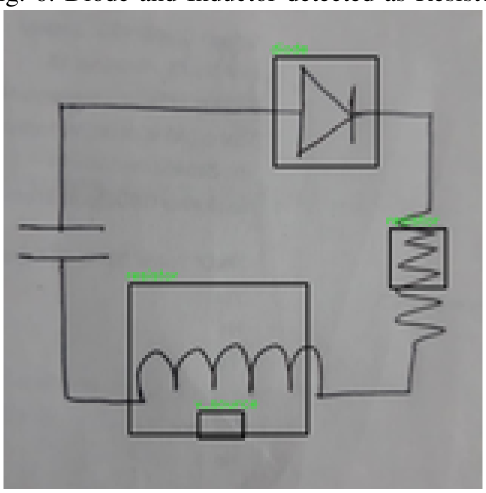


Fig. 7: Inductor detected as Resistor and Voltage Source

C. *Performance Measure*

TABLE I: Performance measure of component detection

	Resistor	Inductor	Diode	Capacitor	Voltage Source	Ground	Not Dectected
Resistor	90	5	0	0	0	5	10
Inductor	10	90	0	0	0	0	10
Diode	5	0	92	0	0	3	8
Capacitor	0	0	0	96	4	0	4
Voltage Source	0	0	0	5	95	0	5
Ground	5	0	0	0	0	95	5

The Fig 3.5 describes the table for the percentage of performance of the recognised components as we have used training and testing ratio as 80 is to 20, i.e, 80- training data and 20-testing data which gave 95% accuracy for Resistors 90% for Inductors and 100% for Diode,90% for capacitor, 95%for voltage source and 95% for ground as well.

IV. Reference Papers

[1] B. Edwards and V. Chandran, “Machine Recognition of Hand-Drawn Circuit Diagrams”, Research Concentration in Speech, Audio and Video Technology, School of Electrical and Electronic Systems Engineering Queensland University of Technology, Brisbane, Australia.

[2] Mayuri D. Patare (Post graduate student) and Madhuri S. Joshi (Professor of CSE Department), “Hand-drawn Digital Logic Circuit Component Recognition using SVM”, Jawaharlal Nehru Engineering college, Aurangabad, Maharashtra, India.

[3] Williams, Kyle; Filho, Milton Ribeiro; and Renshaw, Megan, ”Automatic contextual recognition of hand-drawn content”, Technical Disclosure Commons, March, 2018.

[4] Srikanth Sridar and Krishnan Subramanian, Circuit Recognition Using Netlist Proceedings of the 2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013).

[5] Momina Moetesum1, Syed Waqar Younus1, Muhammad Ali Warsi1 and Imran Sid-diqi1. Segmentation and Recognition of Electronic Components in Hand-Drawn Circuit Diagrams. Research Article: EAI.EU, 10.4108/eai.13-4-2018.154478.

[6] Yasser Alginahi (2010), ”Preprocessing Techniques in Character Recognition”, Character Recognition, Minoru Mori (Ed.), ISBN: 978-953-307-105-3, Talibah University, Kingdom of Saudi Arabia.

Acknowledgement

We take this opportunity to express our profound gratitude and deep regards to our guide **Dr. Leena Ragha** and co-guide **Ms. Dhanashri Bhosale** for their exemplary guidance, monitoring and constant encouragement throughout the completion of this report. We are truly grateful to their efforts to improve our understanding towards various concepts and technical skills required in our project. The blessing, help and guidance given by them time to time shall carry us a long way in the journey of life on which we are about to embark.

We take this privilege to express our sincere thanks to **Dr. Mukesh D. Patil**, Principal, RAIT for providing the much necessary facilities. We are also thankful to **Dr. Leena Ragha**, Head of Department of Computer Engineering, Project Co-ordinator **Mrs. Smita Bharne** and Project Co-coordinator **Mrs. Bhavana Alte**, Department of Computer Engineering, RAIT, Nerul Navi Mumbai for their generous support.

Last but not the least we would also like to thank all those who have directly or indirectly helped us in completion of this thesis.

Ms. Shristi Mittal
Ms. Rhutuja Satpute
Mr. Shubhamm Mohitte