Intelligent Character Reader (ICR) using Artificial Neural Networks (ANN)

Presented By :

Dipayan Deb (13000113032)
Drona Banerjee (13000113033)
Karan Jaiswal (13000113039)
Payal Kumari (13000113060)

Under the guidance of: Prof. Poulami Dutta

Content

Introduction

- conversion of hand-written or printed text to a digital format.
- edit and search the words in the scanned documents like PDF files.

Problem Definition

- Humans identify real life objects easily
- Tough for computers to match human accuracy level
- Train neural network to infer rules for recognition

Uses :

- Used as a form of data entry from printed documents

  Eg: Passport documents, Bank statements, invoices, etc.
- Vehicle license plate identification
- Can be used to read foreign printed languages
- Can be used to arrange letters in post offices
- Helps in digitizing printed document

Benefits: Search
Compact storage
Text Mining
Machine translation

Project Planning :

1. life cycle model
   -Iterative waterfall.

2. Cost analysis

Requirement analysis :

1. Functional requirements
2. Non Functional Requirements
    - Accuracy
    - Data Set for efficient training

The Process :

Image Segmentation → Image Processing → Machine Learning Module

Segmentation Process :

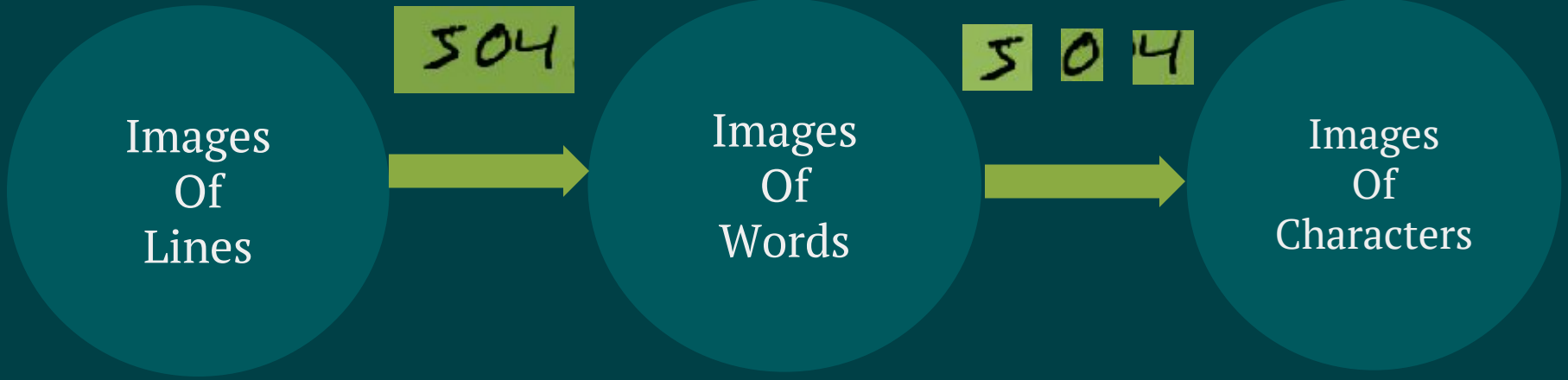
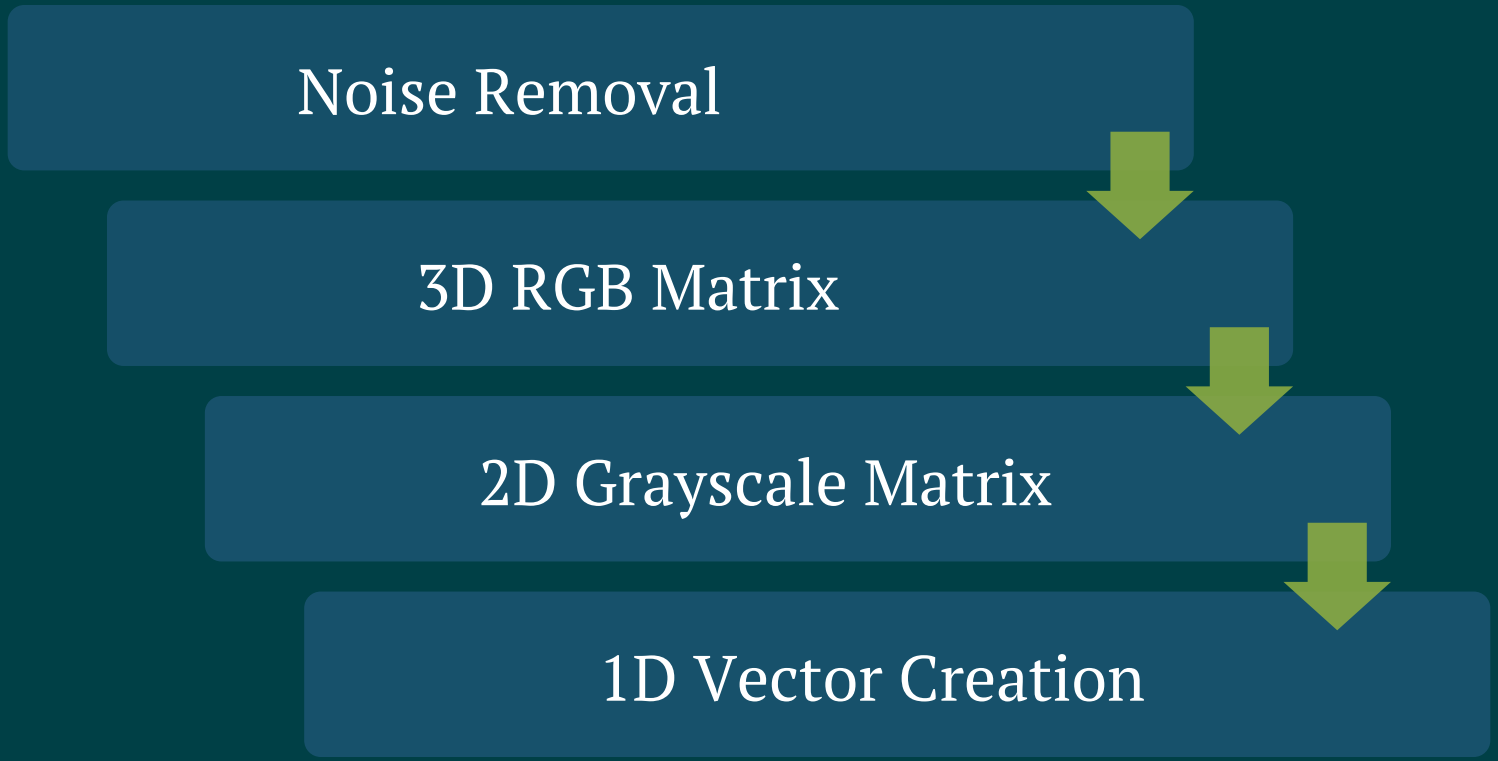


Image Processing :



=

0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	.6	.8	0	0	0	0	0	0
0	0	0	0	0	0	.7	1	0	0	0	0	0	0
0	0	0	0	0	0	.7	1	0	0	0	0	0	0
0	0	0	0	0	0	.5	1	.4	0	0	0	0	0
0	0	0	0	0	0	0	1	.4	0	0	0	0	0
0	0	0	0	0	0	0	1	.4	0	0	0	0	0
0	0	0	0	0	0	0	1	.7	0	0	0	0	0
0	0	0	0	0	0	0	1	1	0	0	0	0	0
0	0	0	0	0	0	0	.9	1	.1	0	0	0	0
0	0	0	0	0	0	0	.3	1	.1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0

Machine Learning :

1. Supervised Learning
2. Classification problem
3. Hypothesis
4. Cost Function
5. Gradient Descent.

Machine Learning : Implementing

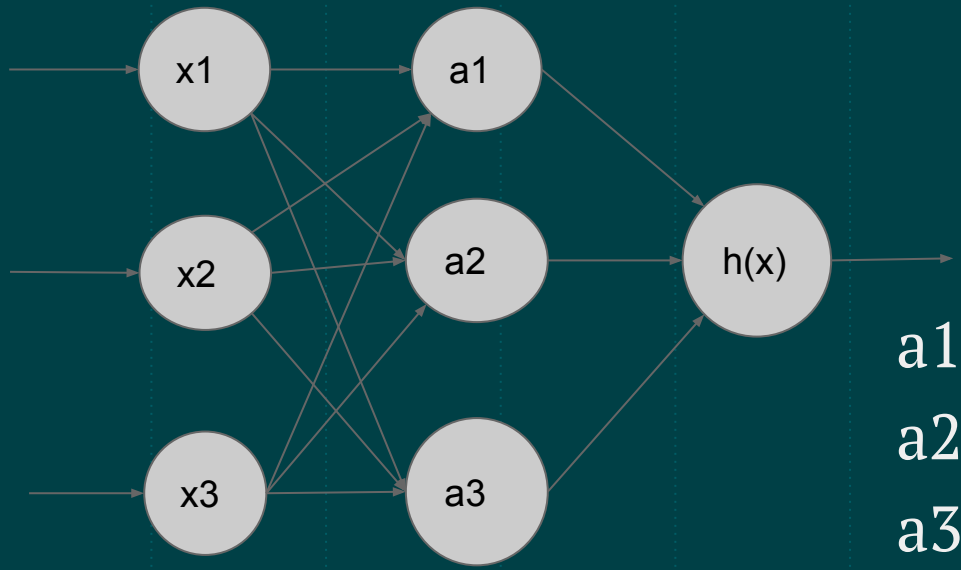
1. Linear Regression

   $$h(x)=g(1+p_1 * \theta_1+p_2 *\theta_2+.....+p_{784} *\theta_{784})$$

2. Neural Network

   -When linear regression (Basic Neural Network) fails.

Neural Network :



$$a1=g(\theta(11)*x1+\theta(12)*x2+\theta(13)*x3)$$

$$a2=g(\theta(21)*x1+\theta(22)*x2+\theta(23)*x3)$$

$$a3=g(\theta(31)*x1+\theta(32)*x2+\theta(33)*x3)$$

$$h(x)=g(\theta(11)*a1+\theta(12)*a2+\theta(13)*a3)$$

Machine Learning Module :

Number Module

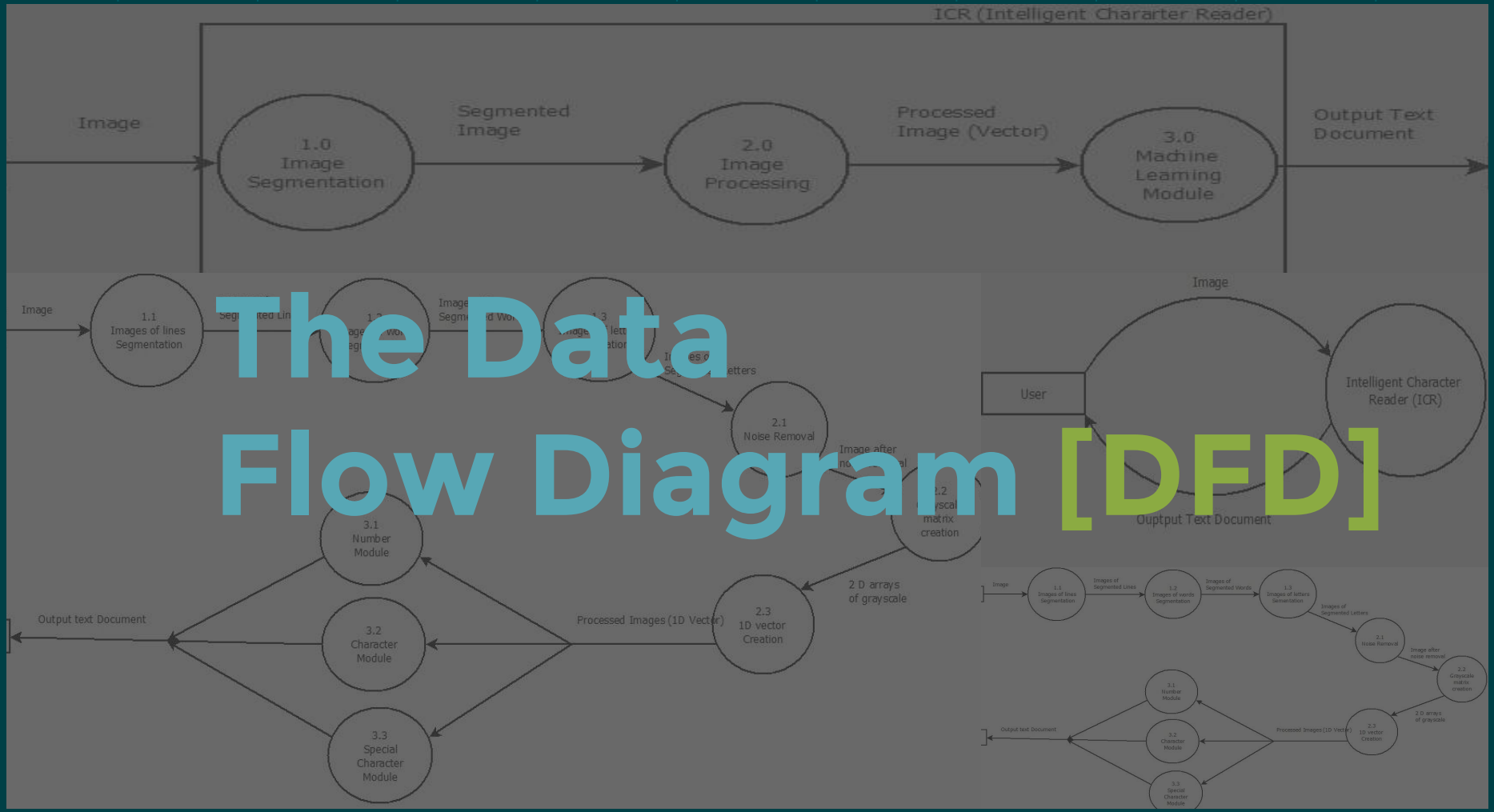
- Logistic Regression
- Classification Problem
- Sigmoid Function

Character Module

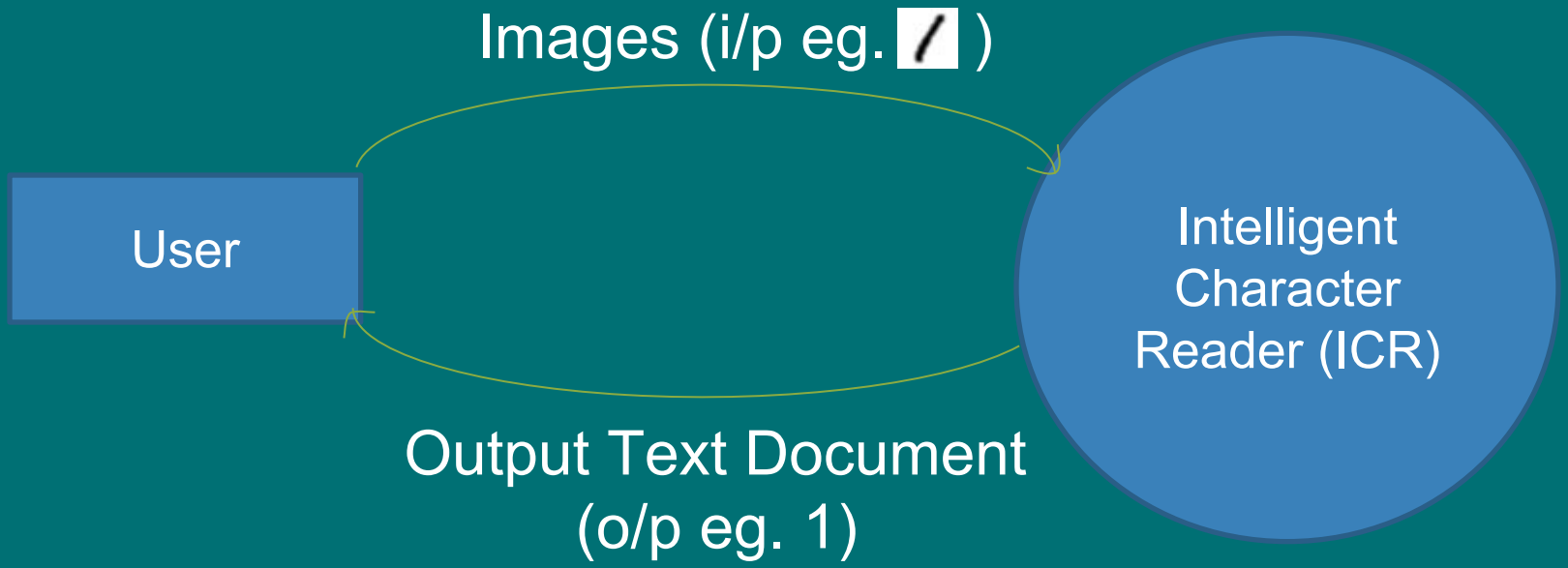
- Neural Networks
- Input : Pixel intensities of 1D array
- 3 Layer Neural Network

Special Character Module

- Logistic regression
- Classification Problem
- Sigmoid Function



The Data Flow Diagram [DFD]



(a) Level 0 (Context Diagram)

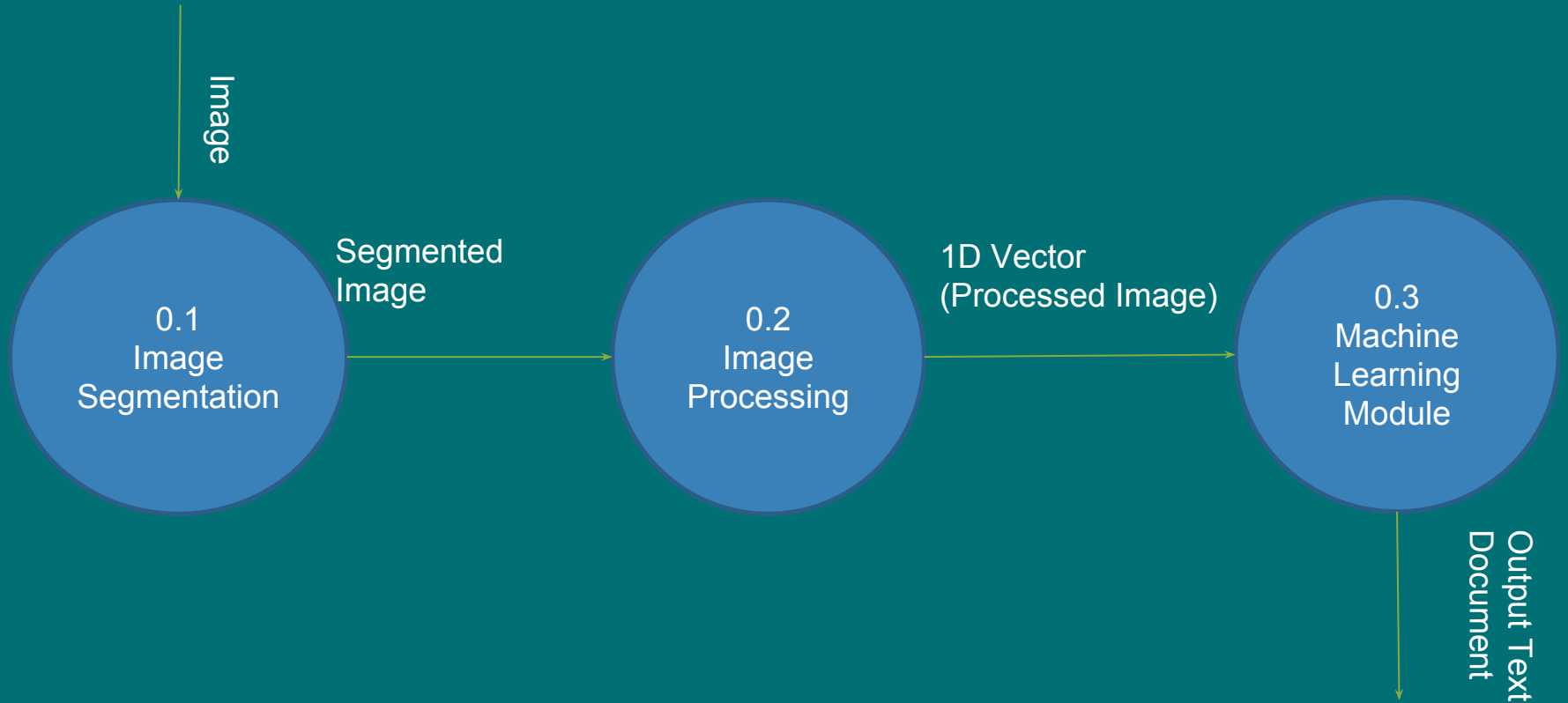
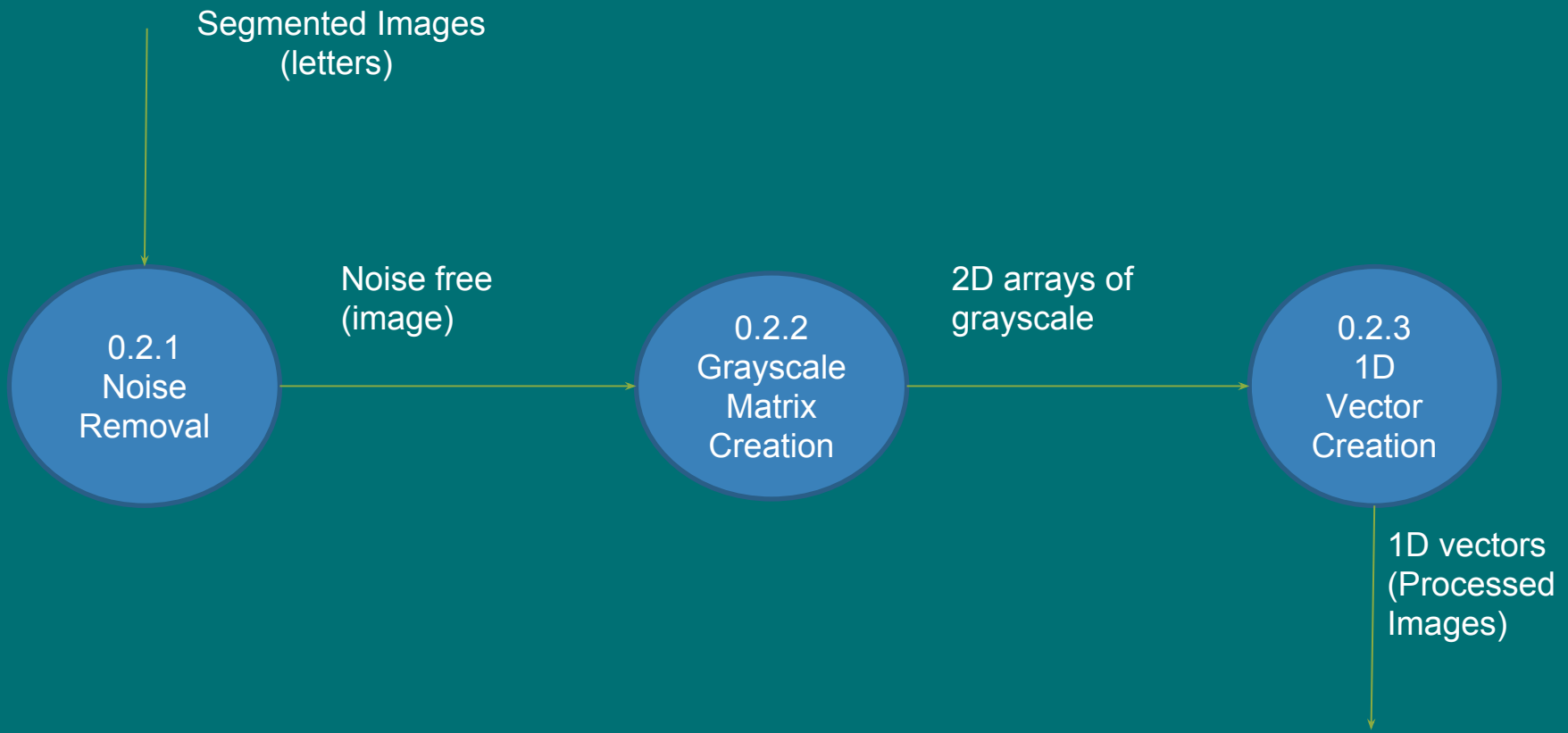


Image
0.1
Image
Segmentation
Segmented
Image
0.2
Image
Processing
1D Vector
(Processed Image)
0.3
Machine
Learning
Module
Output Text
Document

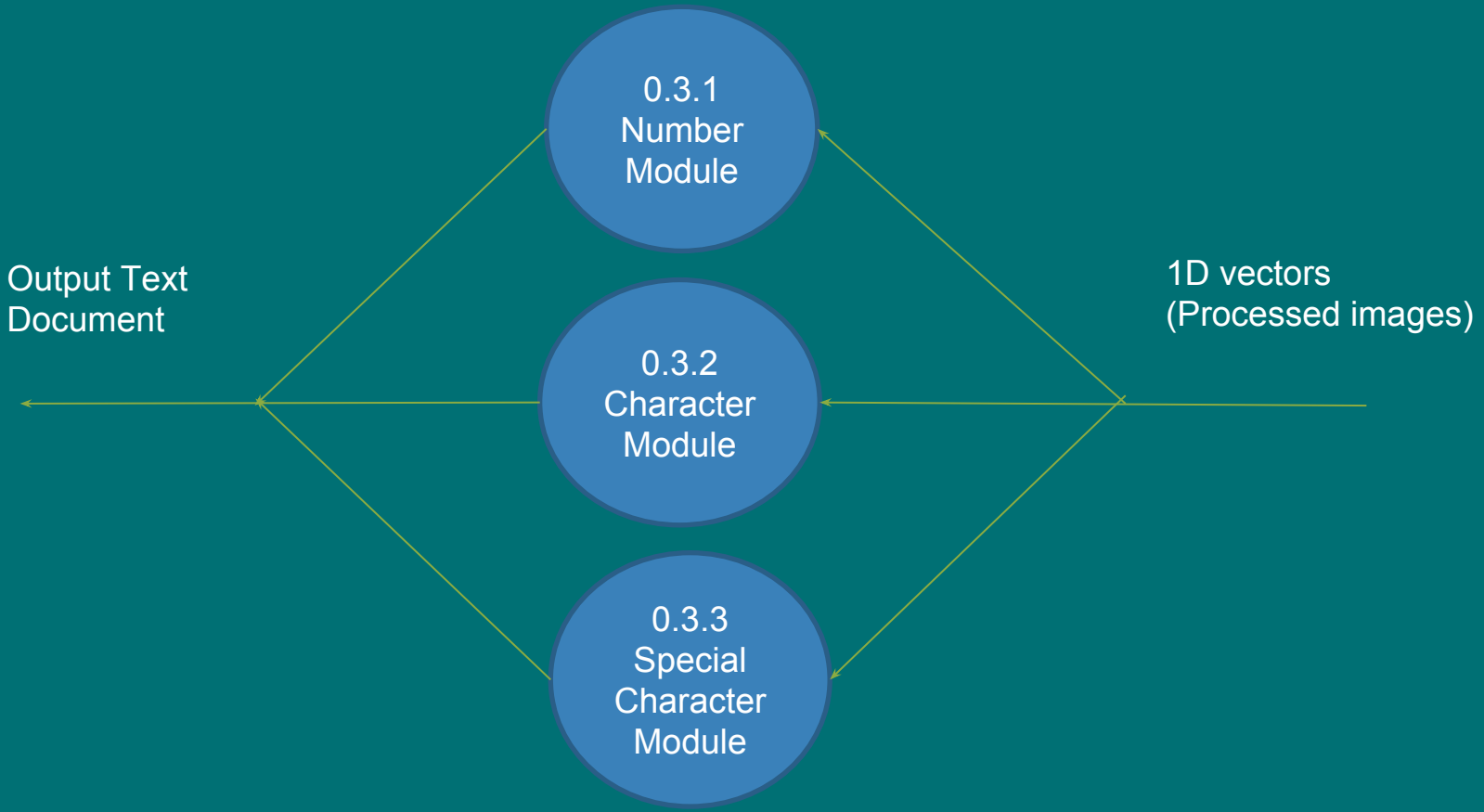
(b) Level 1


Image
0.1.1
Images of
Segmented
lines
Segmented image
(lines)
0.1.2
Images of
Segmented
words
Segmented image
(words)
0.1.3
Images of
Segmented
letters
Segmented image
(letters)

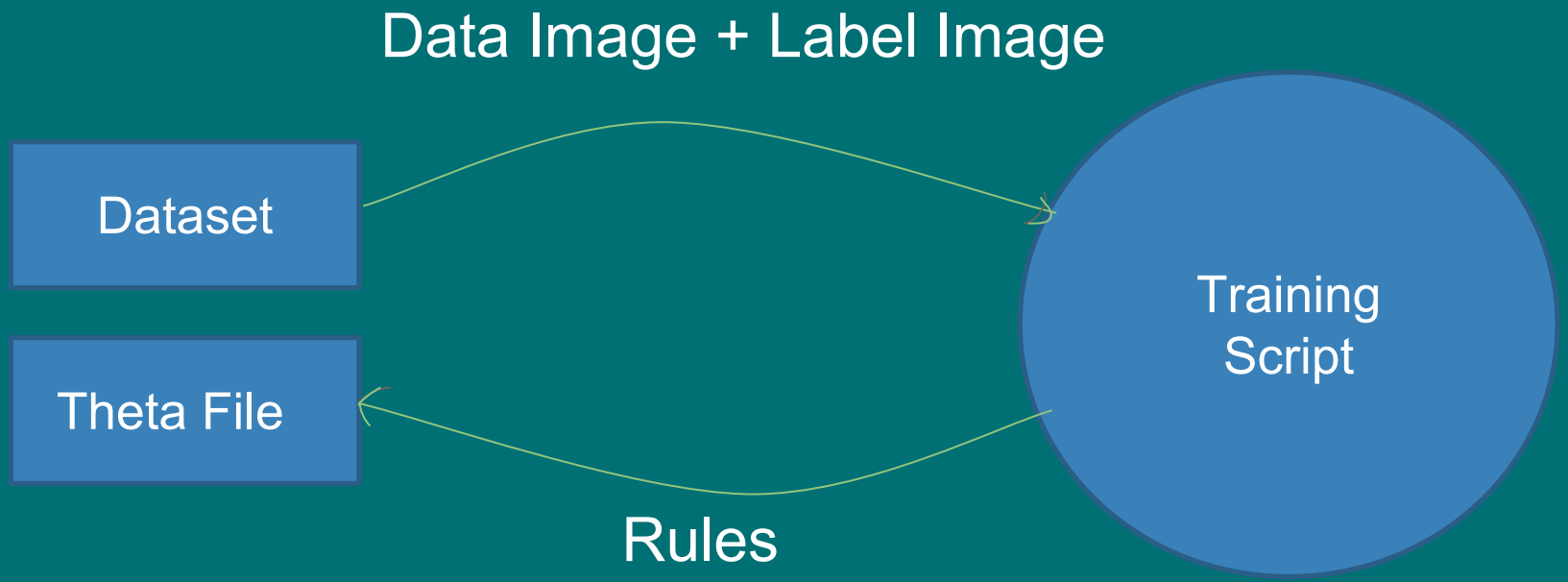
(c) Level 2


Segmented Images
(letters)
0.2.1
Noise
Removal
Noise free
(image)
0.2.2
Grayscale
Matrix
Creation
2D arrays of
grayscale
0.2.3
1D
Vector
Creation
1D vectors
(Processed
Images)

(c) Level 2


0.3.1
Number
Module
0.3.2
Character
Module
0.3.3
Special
Character
Module
Output Text
Document
1D vectors
(Processed images)

(c) Level 2



(a) Level 0 (Context Diagram for Training Script)

Analysis :

Strength	Weakness	Opportunities	Threats
• Offline Functioning • Automation Using AI • Omni-font Featurization	• High Hardware Requirements • Difference In Calligraphy (Style & Direction)	• Expanding Digitization Market • Modular Extension Of Other Languages	• Absence Of Preset Database For Training • No Benchmark Set For Unconfigured Languages

Related Study :

TESSERACT OCR

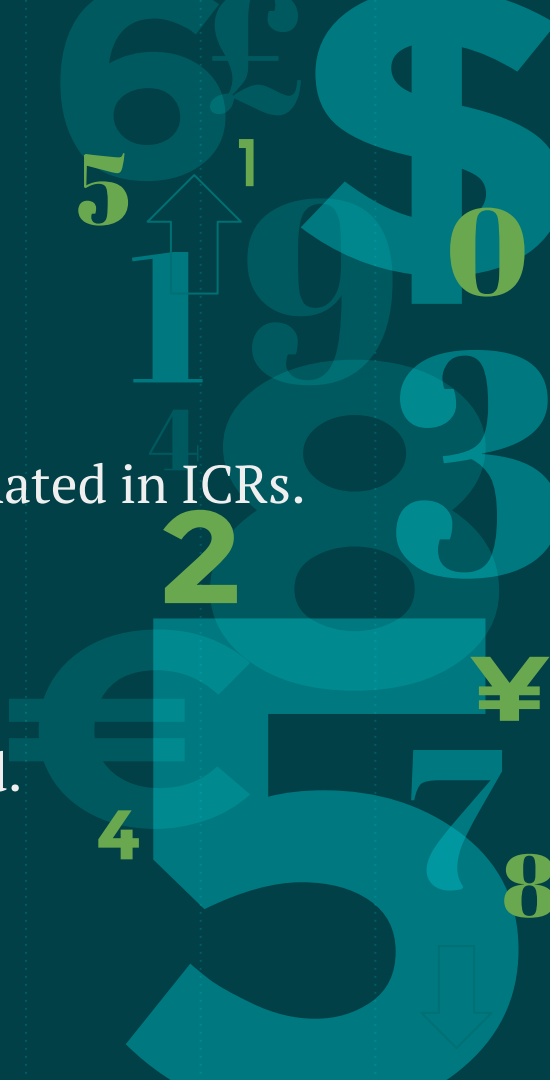
- Not an application, an engine
- 3rd party software required
- Does not focus on being user-friendly

CUNEIFORM OCR

- Does not recognise handwritten text
- Does not recognise decorative fonts
- Accuracy drops when source image quality is low

Conclusion

- Use of Neural Networks give high accuracy rate.
- Human intervention needed in OCR’s can be eliminated in ICRs.
- Eg. Letter sorting machines
- Useful for automated form processing.
- Can be helpful for people with eye problems to read.



Future Scope of improvement :

> Intelligent Word Readers can be built that process one word at a time.

> Interpreting the sentence on basis of prior knowledge and understanding can be implemented.

> Using Natural Language Processing , Artificial Intelligence and incredibly advanced machine learning.

>Modules of any other language can be added.

References :

- https://www.coursera.org/learn/machine-learning
- https://en.wikipedia.org/wiki/Machine_learning
- http://ieeexplore.ieee.org/document/1688109/

Thank You!