Flowchart Recognition

from Hand-drawn Flowchart to Digital Components

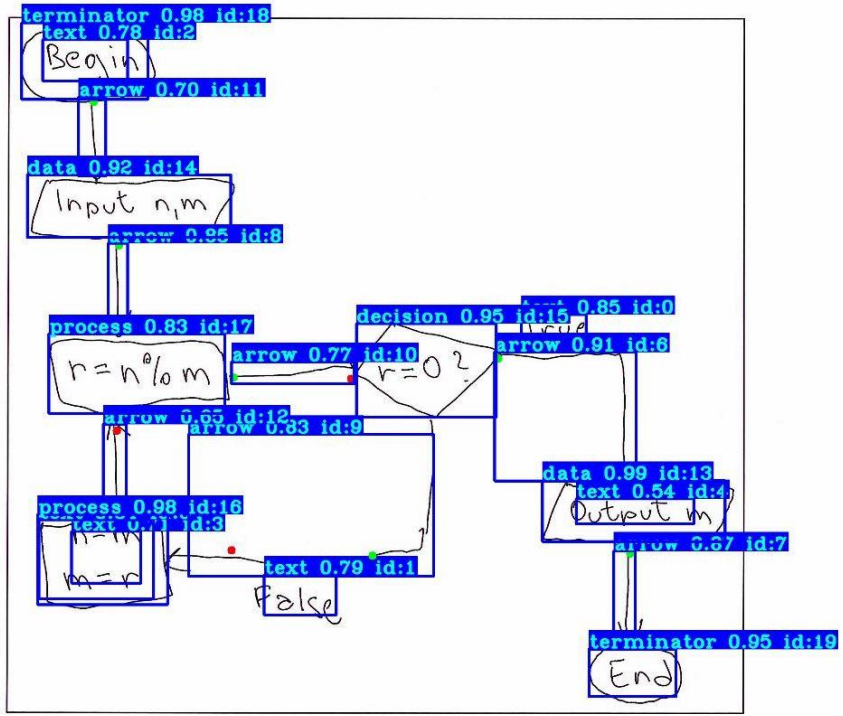
SJSU CMPE 258

Mu Chen, MS-SE (mu.chen@sjsu.edu) [SJSU ID: 014725425]
Roger Kuo, MS-SE (roger.kuo@sjsu.edu) [SJSU ID: 013784706]
Hardy Leung, MS-AI (kwok-shing.leung@sjsu.edu) [SJSU ID: 016711877]
Jasmine Wang, MS-AI (jasmine.wang@sjsu.edu) [SJSU ID: 002805362]

Why Flowcharts?

- Incredibly versatile and flexible
  - Succinctly summarize or visualize concepts
- PPT, Keynote, Google Docs all support FC
  - Very primitive
  - Hard to use
  - Mismatch between intents and outcomes
  - Disappointment
- **Deep Learning to the rescue**
  - translate user's intent to
  - accurate rendering of flow chart

Architecture Overview



- Custom-trained YOLO engine
- Processing + Transformation + Analysis
  - Connectivity
  - Symbolic Relations
  - Text OCR
- Result = Symbolic Flowchart
  - Intents fully captured
  - Can be rendered in several ways

Dataset and Preprocessing

- 600+ high-quality flowcharts
  - With xml annotations
- 80/20 training/testing split
- Annotation conversion
  - From xml to YOLO format

Training

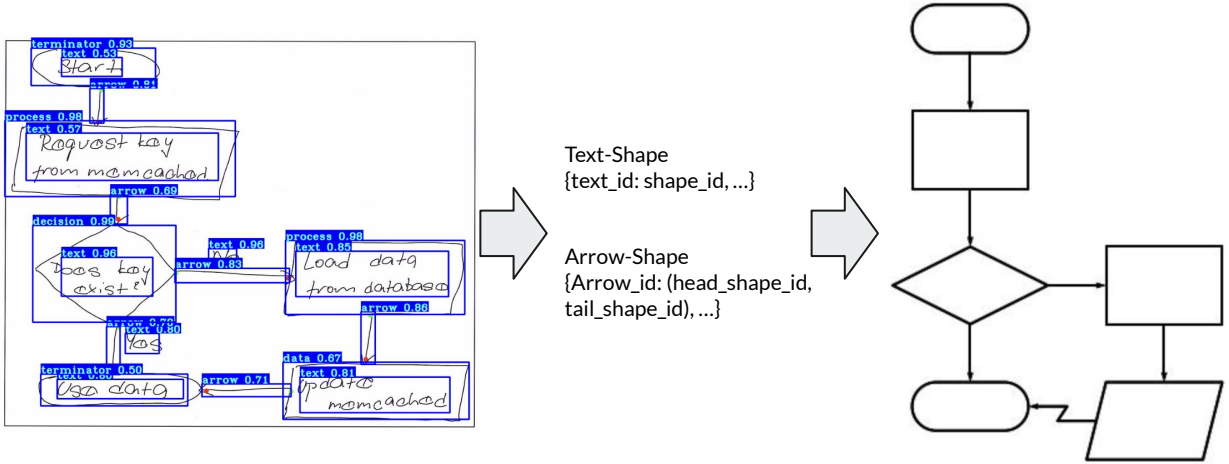
- Custom YOLOv4 Model
  - Tiny YOLOv4 weights
- Seven custom classes
  - Text; Arrow; Data; Decision, Process; Terminator; Connection
- Achieved 90%+ mAP
  - 96%+ accuracy on all flowchart elements except arrows and texts

Detection

- List objects with attributes
  - Position, Class, etc
- Arrow head/tail detection
  - K-means clustering

Flowchart Rendering

- With all recognized objects
- Extract arrow to shape relationships
- Extract text to shape relationships
- SchemDraw
  - Display flow chart based on shape relationships
  - Drawback: Rely heavily on the correctness of detection
- Tesseract for text recognition (need improvement)



Discussion

- Results:
  - Achieved 90%+ mAP with 96%+ accuracy on all flowchart elements except arrow and text
- Challenges:
  - Arrow and text detection needs improvement
    - Current YOLO precision too low
  - Arrow head/tail detection needs improvement
    - Current method not resilient to noise
  - Schemdraw is not the most optimal rendering tool
    - Current method heavily relies on detection accuracy
  - Tesseract needs to be trained specifically for handwritten text
- Future Plans:
  - Larger dataset to cover more variety and migrating to HPC to accelerate training
  - Overall code efficiency improvement including semi-automatic testing

References

- Python scripts

  https://pylessons.com/

- Yolo v4

  https://github.com/pythonlessons/TensorFlow-2.x-YOLOv3