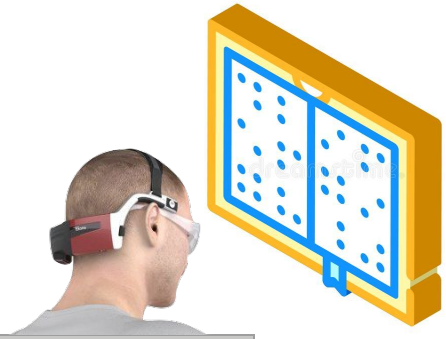
Carnegie Mellon University

Aware-ables

Team B1: Chester Glenn, Jong Woo Ha, Kevin Xie
Presented by Jong Woo (Jay) Ha

Use Case and Requirements



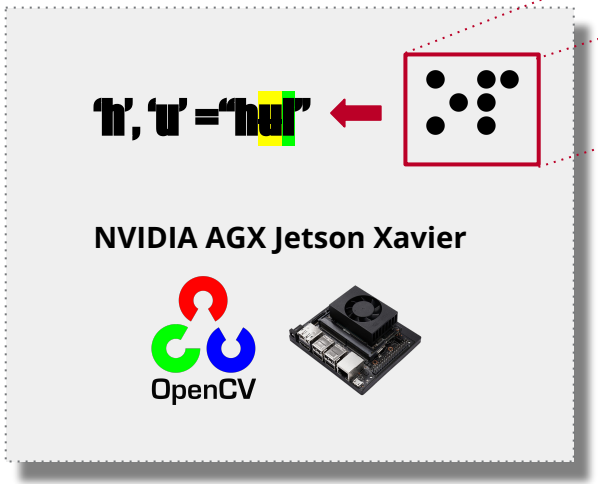
"A new device for auditory accessibility and assistance"
Wearable braille detection for increased awareness of surroundings and braille literacy

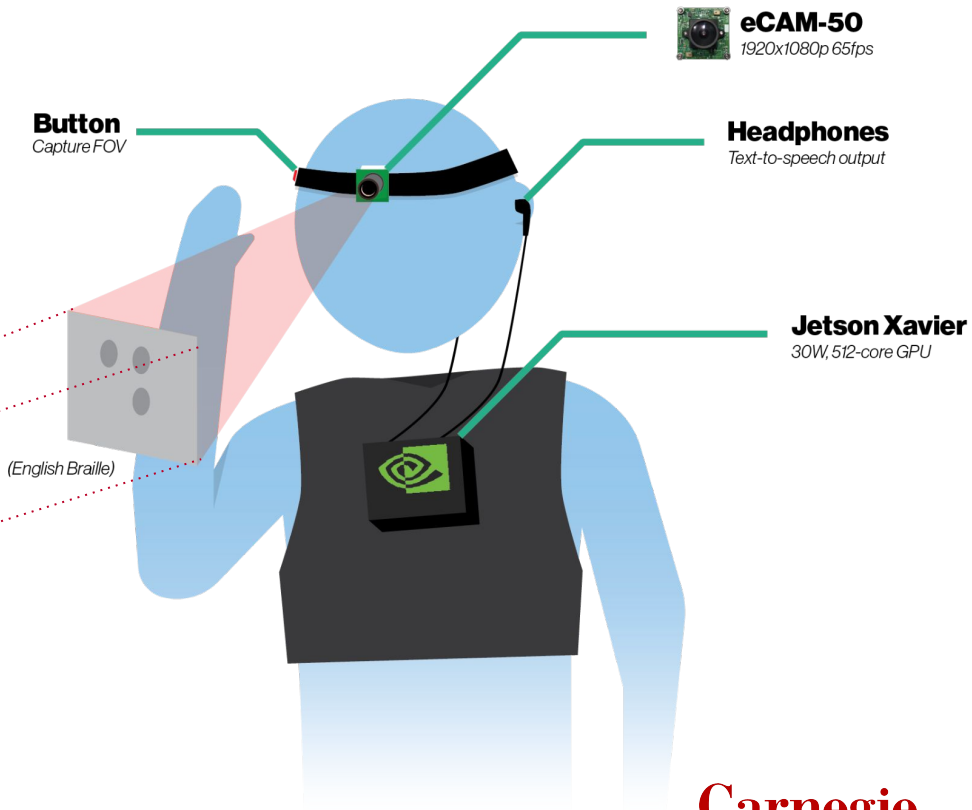
Requirement	Target (units)	Justification
Text-to-Speech latency	2s	Common usability standard for loading wait times
Words per Frame	>10	~300wpm to match braille reading speed *NOTE: 150wpm is a comfortable speaking speed
Character Error Rate	10%	Matches error rate of traditional OCR
Word Error Rate	<10%	Errors should be corrected from spellcheck

Carnegie Mellon University

Solution Approach

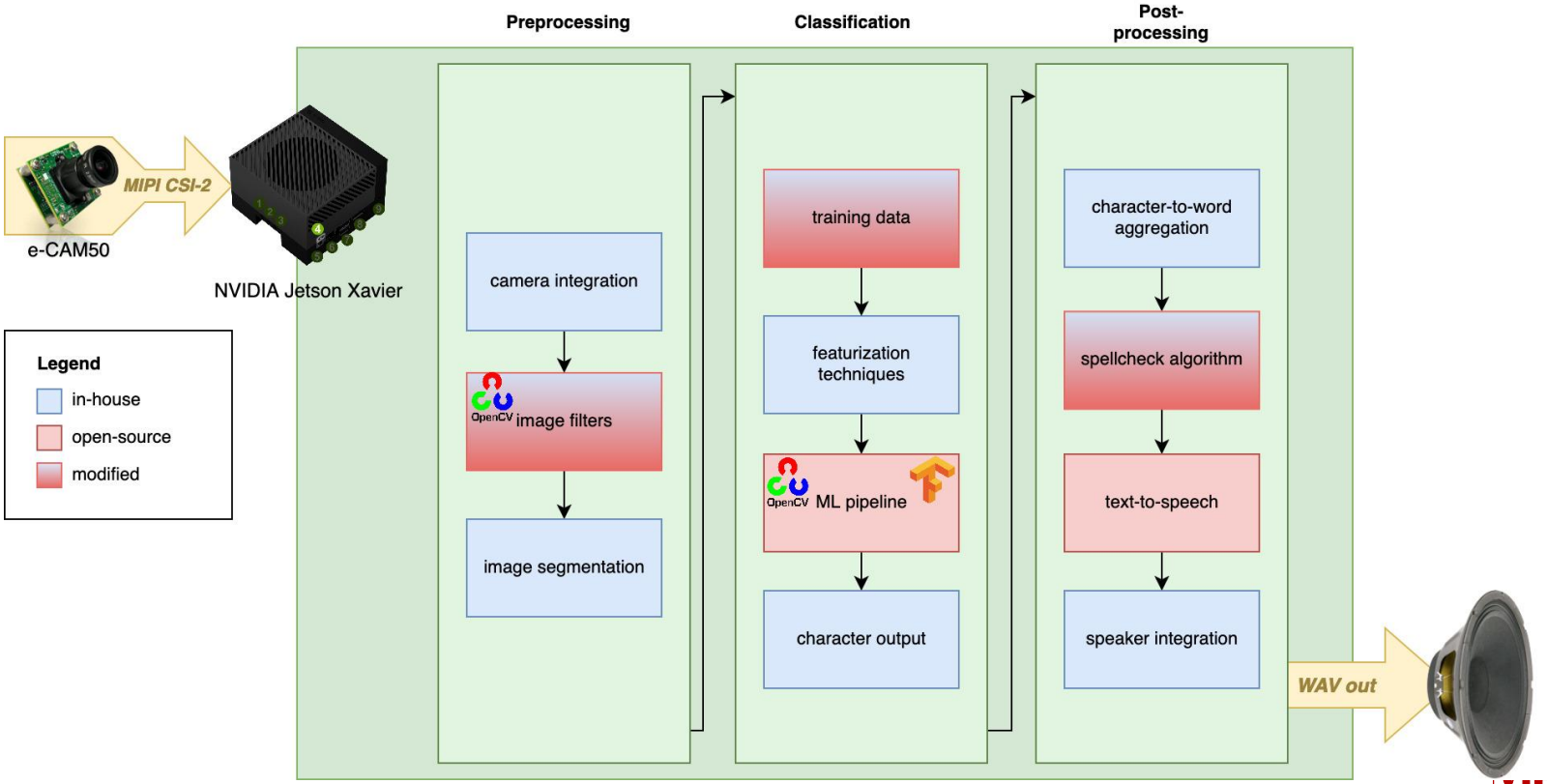
- Head-level camera for pointing at Braille writing
- Button for easily triggering Braille capture
- Jetson AGX Xavier vs. Jetson Nano



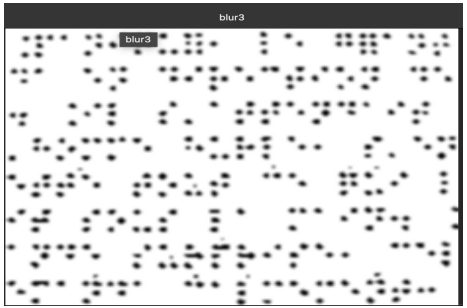


Carnegie Mellon University

Block Diagram



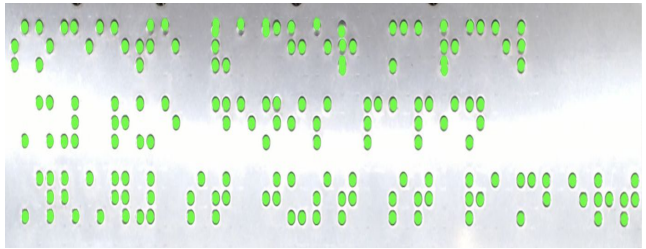
Pre-processing

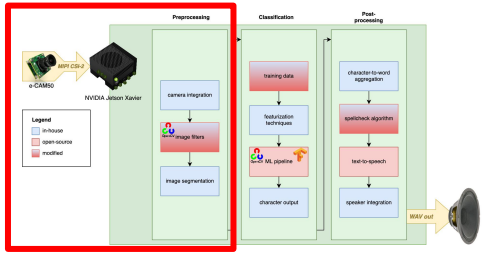


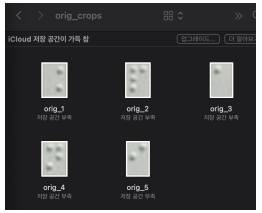
Original braille document image with horizontal segmentation / Pre-processed images after blurring | thresholding

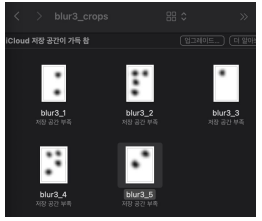
vs

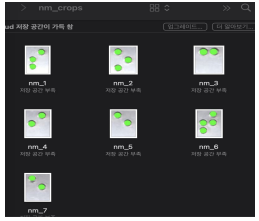
Further pre-processing using non-maximum suppression algorithm











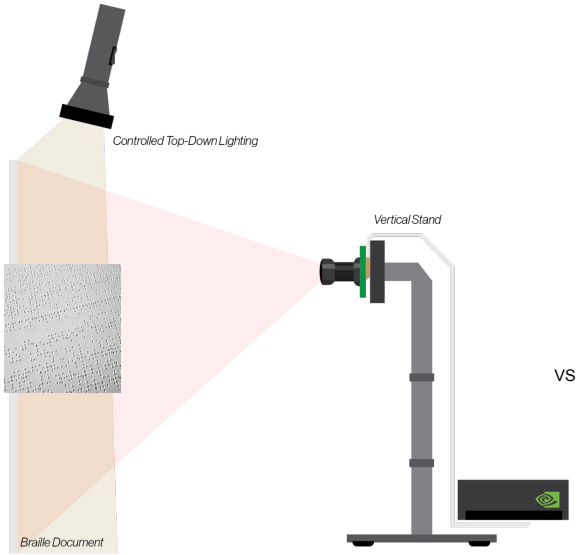
- Primary Goal: **~50ms** inference latency / attain minimum thresholds for classification phase's **~30ms** latency
- Initial cropping -> Pre-processing -> Segmentation ; result: folder of individually cropped brailes
- cv2.cvtColor(src,code[,dst[,dstCn]])
- cv2.blur() / cv2.GaussianBlur() / cv2.medianBlur() / cv.bilateralFilter()
- cv2.threshold(), cv2.adaptiveThreshold()
- non-maximum suppression

Carnegie Mellon University

Pre-processing - Design choices for Initial Cropping given Natural Scene Braille Detection

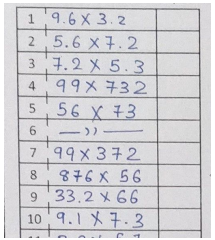
Manual Cropping

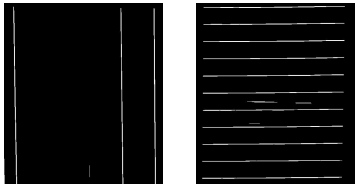
VS



Approach with controlled dimensional guidances

VS

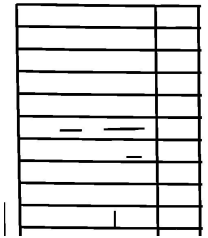


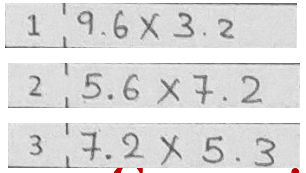


Example of ML based recognition

Images starting from above going clockwise.

1. sample image
2. vertical lines
3. horizontal lines
4. containing only boxes
5. extracted images





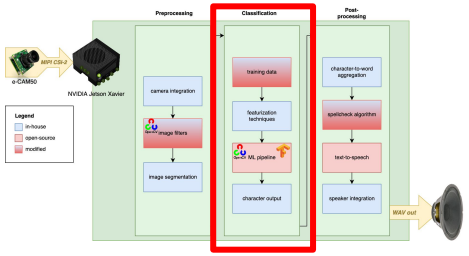
braille document image source: https://www.researchgate.net/profile/Giovanni-Farinella/publication/287222520/figure/fig1/AS:583701360574466@1516176657486/Example-of-Braille-document-Each-character-consists-of-a-maximum-of-six-dots-arranged-in.png

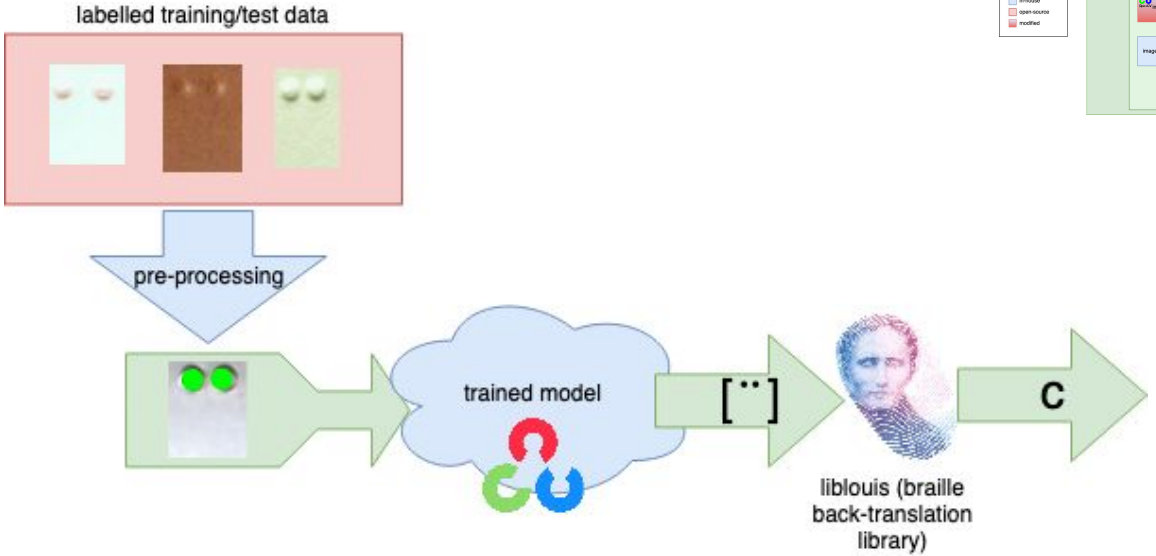
braille document image source: https://upload.wikimedia.org/wikipedia/commons/thumb/e/ee/Braille_text.jpg/749px-Braille_text.jpg

ml based recognition example source: https://medium.com/coinmonks/a-box-detection-algorithm-for-any-image-containing-boxes-756c15d7ed26

Carnegie Mellon University

Classification

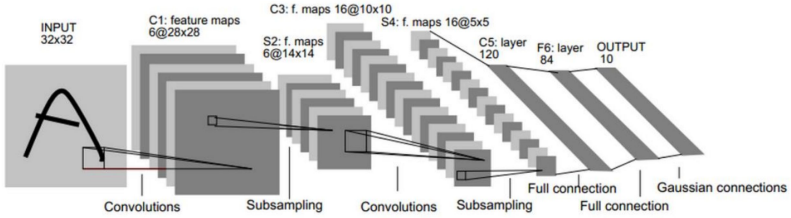




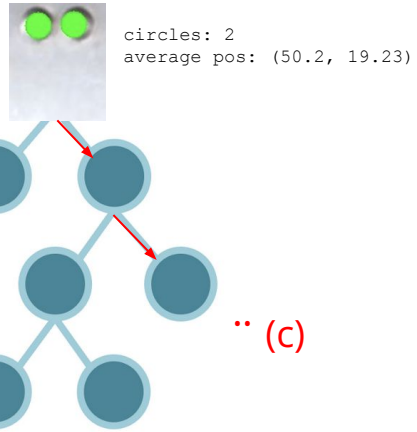
- 30,000 image open-source dataset
- **~30ms** inference latency based on some experiments
- Open-source braille back-translation library for future extensibility

Carnegie Mellon University

Classification: DNN vs. Tree-model



Gradient-based learning applied to document recognition. LeCun et al., 1998



Neural Networks:

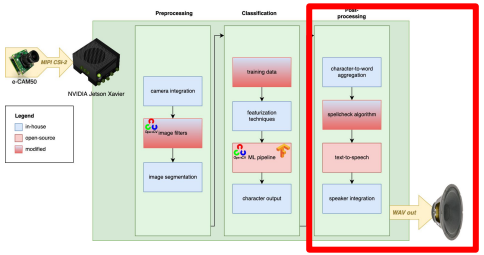
- Good for large datasets
- Much larger training overhead
- Likely better accuracy
- Opaque

Tree-model:

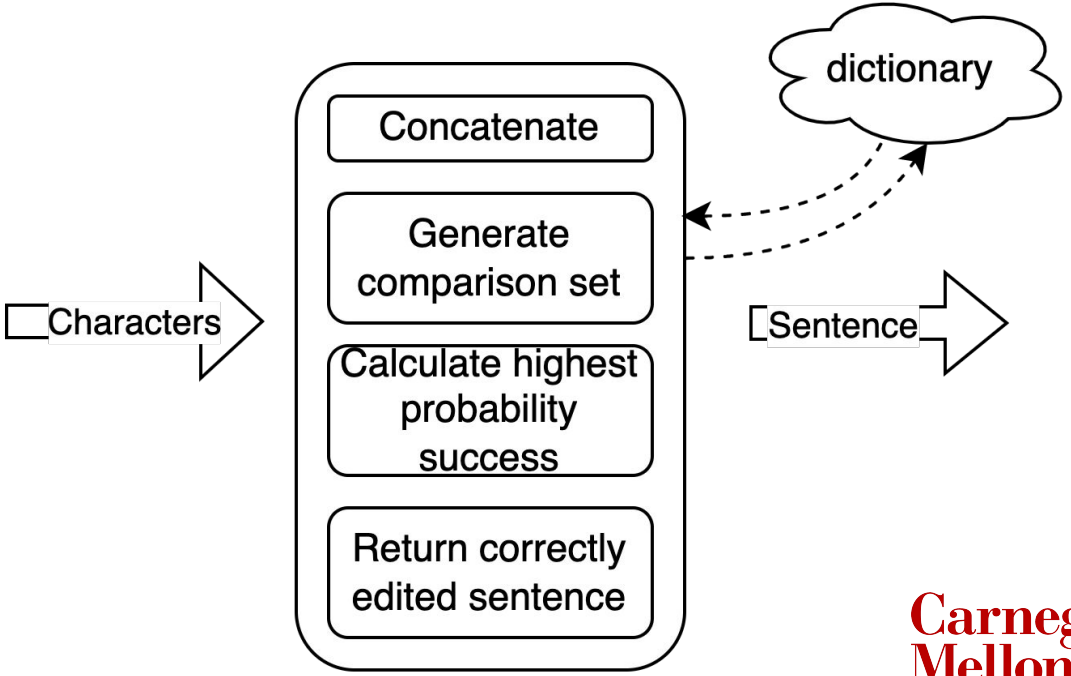
- More controllable feature selection
- More interpretable result
- Likely faster inference

Carnegie Mellon University

Post-Processing: Spellcheck



- Monitor inflow of characters and concatenate to words and sentences.
- Each word will be individually processed
- At the end a sentence will be formed given the corrected words

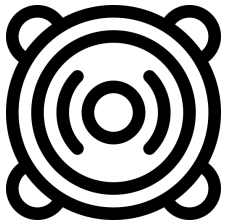


Carnegie Mellon University

Post-Processing: text-to-speech

- Complexity analysis showed that installing a text to speech API would be significantly more efficient and logical for usability
  - Text to speech requires in-depth signal processing and ML for smooth life like speech.

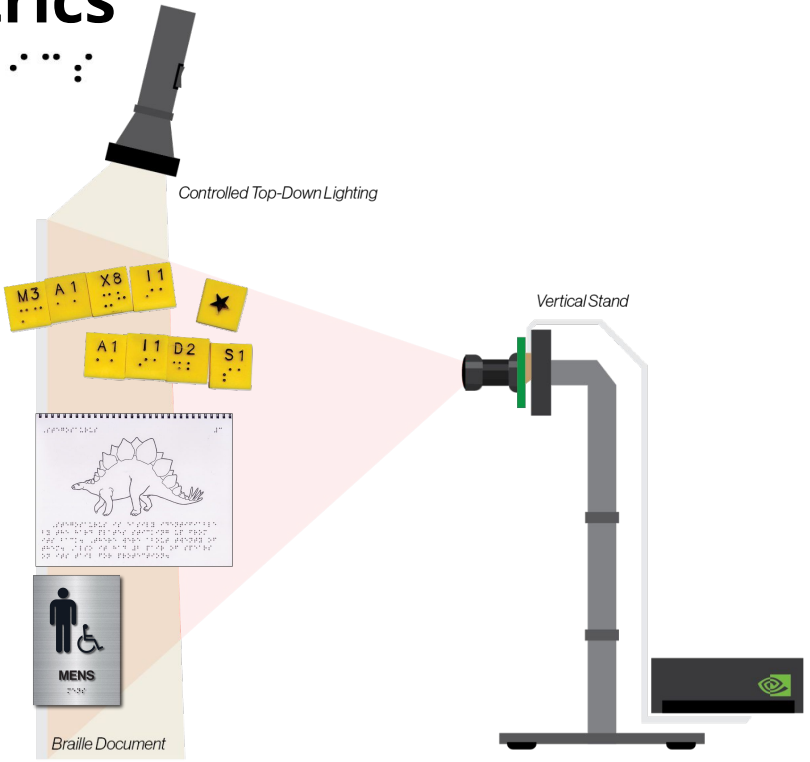
- API will generate WAV files for direct connection to USB speaker (headphones, earbuds, speaker)



Carnegie Mellon University

Testing, Verification, and Metrics

Metric	Test Input	Risk Factors / Unknowns
Character Error Rate (10%)	Braille tiles and open-source dataset	ML model may overfit to training data (would need to retrain/prune model)
Word Error Rate (<10%)	Braille books / signs	Needs to ensure that spellcheck is accurate and accounts for Braille quirks
Latency (2 sec)	Braille books / signs	Will need to tune text-to-speech speed for comfortable comprehension (~150wpm)
Usability (MVP+)	Power efficiency / wearable form factor	Xavier power draw, weight



Carnegie Mellon University

Schedule

Aware-able Schedule

Read-only view, generated on 02 Oct 2022

Instagantt

	ACTIVITIES	ASSIGNEE	EH	START	DUE	%
	Research + Development:		-	14/Sep	21/Sep	100%
1	Proposal Presentation Slides	CG, Ja, KX	-	14/Sep	18/Sep	100%
2	Proposal Presentation	Kevin Xie	-	19/Sep	21/Sep	100%
3	Wordpress Bring-up	Jay	-	17/Sep	17/Sep	100%
	Design Phase:		-	22/Sep	10/Oct	58%
5	Hardware Design	CG, Ja, KX	-	22/Sep	26/Sep	100%
6	Finalize Parts List		-	04/Oct	07/Oct	0%
7	Order Parts		-	10/Oct	10/Oct	0%
8	Software Architecture Design	CG, Ja, KX	-	22/Sep	28/Sep	100%
9	Design Presentation Slides		-	29/Sep	03/Oct	100%
10	Design Presentation		-	03/Oct	05/Oct	0%
	Development Phase:		-	03/Oct	25/Nov	0%
12	Interim Demo	CG, Ja, KX	-	07/Nov	07/Nov	0%
	Hardware Bring-up		-	03/Oct	18/Nov	0%
14	Camera Integration	Jay	-	03/Oct	11/Oct	0%
15	Speaker Integration	Chester Glenn	-	09/Nov	18/Nov	0%
	Image preprocessing and s...	Jay	-	03/Oct	22/Nov	0%
17	Research on braille det...		-	03/Oct	10/Oct	0%
18	Design image pre-proc...		-	11/Oct	26/Oct	0%
19	Design segmentation al...		-	27/Oct	10/Nov	0%
20	Validation of approach ...		-	11/Nov	22/Nov	0%
	Character recognition and c...	Kevin Xie	-	03/Oct	25/Nov	0%
22	Research/Test Existing ...	Kevin Xie	-	03/Oct	11/Oct	0%
23	Build dataset	Kevin Xie	-	12/Oct	19/Oct	0%
24	Develop & Train CV Mo...	Kevin Xie	-	20/Oct	04/Nov	0%
25	Validate CV	Kevin Xie	-	08/Nov	11/Nov	0%
26	CV Revisions	Kevin Xie	-	21/Nov	25/Nov	0%
	Postprocessing & Text-to-sp...	Chester Glenn	-	03/Oct	18/Nov	0%
28	Design post-processing...	Chester Glenn	-	03/Oct	11/Oct	0%
29	Maintain buffer of char...	Chester Glenn	-	12/Oct	20/Oct	0%
30	Concatenate to words	Chester Glenn	-	21/Oct	31/Oct	0%
31	Process words & check ...	Chester Glenn	-	25/Oct	08/Nov	0%
32	Final text to voice integr...	Chester Glenn	-	04/Nov	18/Nov	0%
	Unit Testing	CG, Ja, KX	-	03/Oct	25/Nov	0%
34	Character recognition l...	Kevin Xie	-	14/Nov	18/Nov	0%
35	Slack Time		-			0%
	Testing & Revision Phase:		-	21/Nov	02/Dec	0%
37	Final Integration Testing		-	21/Nov	02/Dec	0%
38	Revisions		-	21/Nov	02/Dec	0%
	Wrap-up:		-	28/Nov	05/Dec	0%
40	Final Presentation Slides		-	28/Nov	05/Dec	0%
41	Final Presentation		-	05/Dec	05/Dec	0%
42	Final Report		-			0%
43	Final Video		-			0%
44	Public Demo		-			0%