Capstone Oral Exam

Requirements

- 20-24 slides
- 40-45 minute time slot
- 10% of final grade
- ***What do we need to include in this??***

Marking Scheme

Problem Definition - 20

Design Choices, Analysis, and Decision - 50

Clarity and Concise Presentation - 15

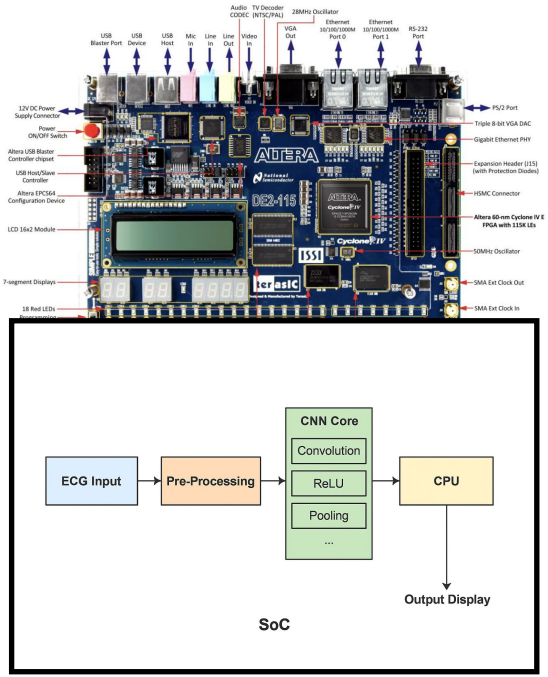
Preparedness for COE 70B (Implementation Plan) - 15

Individually Marked

Malcolm’s 6 Slides (15 minutes)

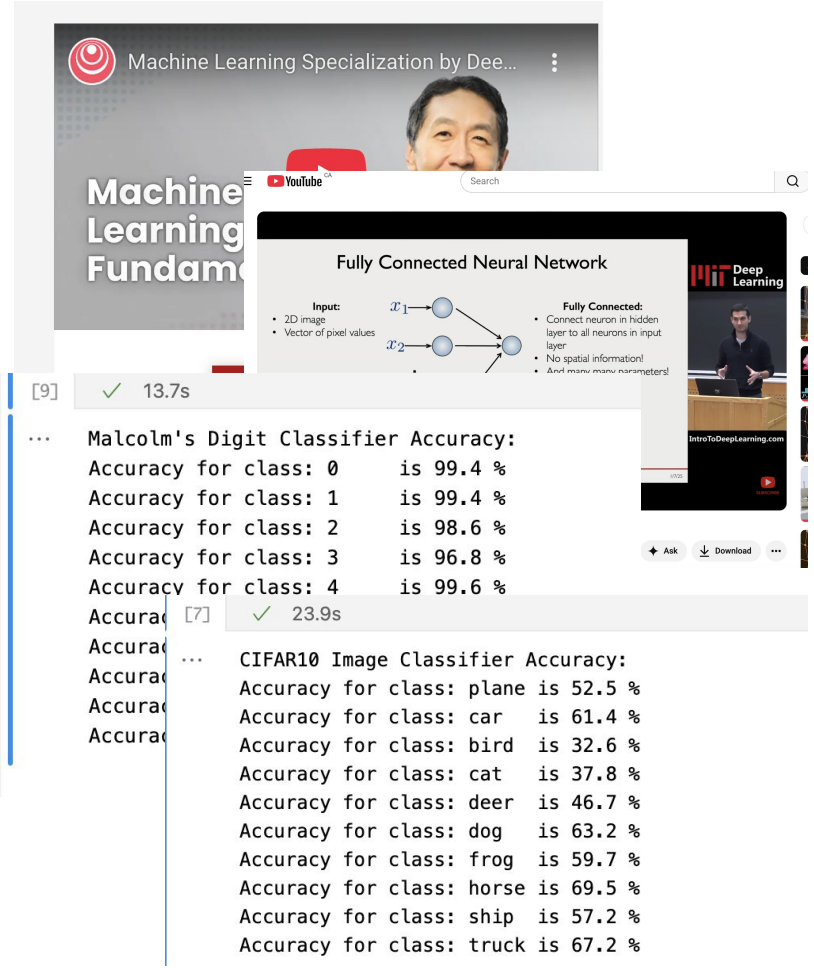
Project Plan

- Divided work into logical, loosely coupled subteams, that can work concurrently.
  - CNN Model → Ayoub; SoC Team → Malcolm, Pierre; Simulation Component → Marly
- CNN Team
  - Deliverables: CNN-based architecture + model weights
- SoC Team
  - Deliverables: All SoC hardware that implements CNN model
- Simulation Component Team
  - Deliverables: simulation device that visually demonstrates our system
  - Implemented using FPGA to facilitate learning.



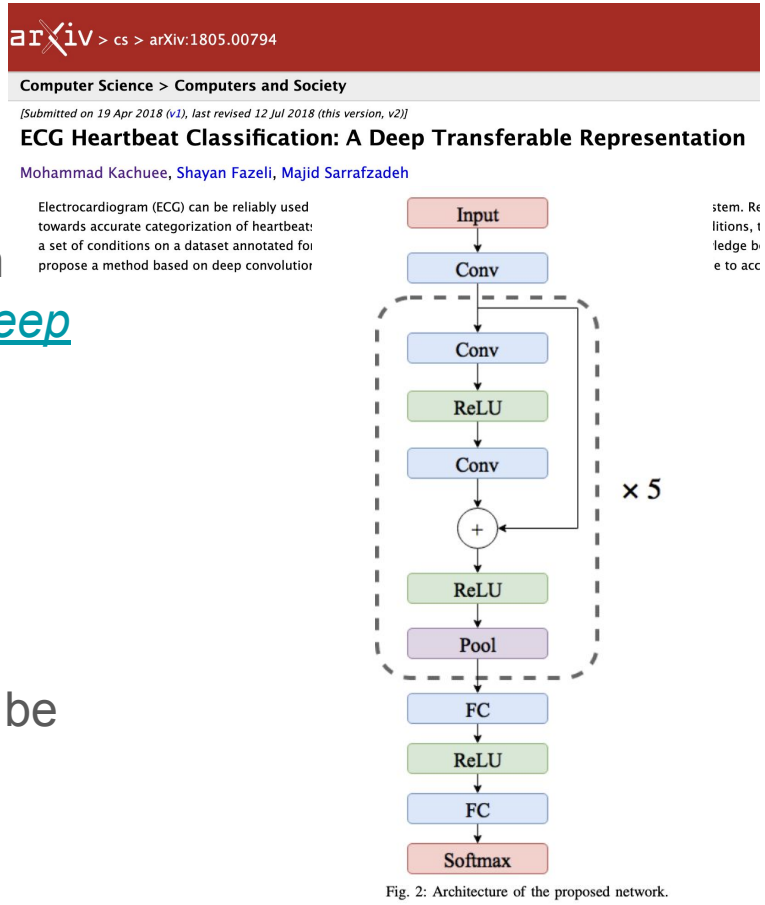
Neural Networks

- Last time I tried to learn about DL was in 2015. I kicked the can down the road after realizing the math that went into it.
- Worked through 6.S191 to understand basic of DL.
- Made some practice models: image classifier, digit classifier



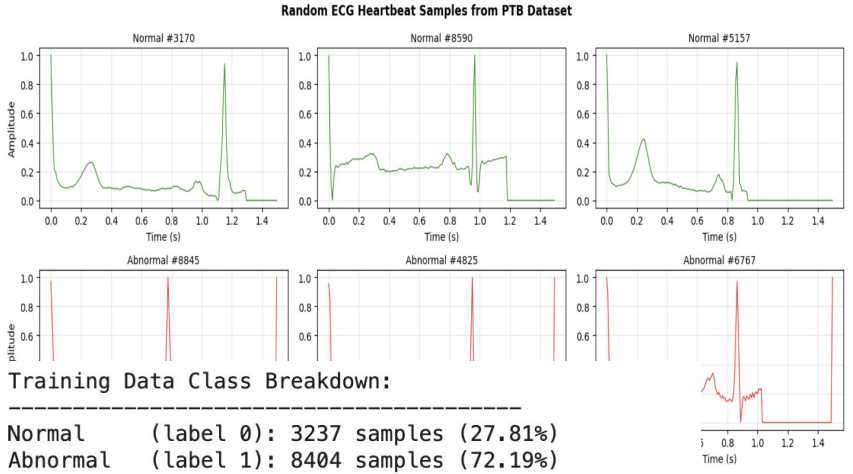
Neural Networks (Model)

- Read through a variety of papers to find a simple CNN-based architecture. Decided on this one: *ECG Heartbeat Classification: A Deep Transferable Representation*
  - 13 weight layers total: 1 initial conv + (5 blocks × 2 convs) + 2 FC = 13
  - Easy to implement in SW for DL noob. Hopefully implementable on hardware!
- Current model hosted here on Colab (might be some bugs!)
  - Attained 88% abnormal recognition, 92% normal recognition. Possibilities to improve here.



Neural Networks (Data)

- Based on the Kaggle heartbeat dataset. Pre-processed data from MIT + PTB ECG datasets.
    - PTB: 125 samples/s, 188 samples each, normal/abnormal categories, single channel, 1 heart beat per csv row. Contains ~500 heartbeats.



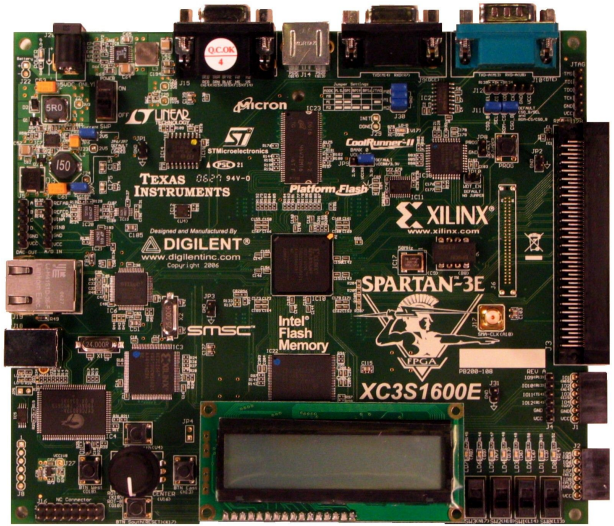

```
Training Data Class Breakdown:
----------------------------------------
Normal     (label 0): 3237 samples (27.81%)
Abnormal   (label 1): 8404 samples (72.19%)
----------------------------------------
Total: 11641 sam
```

```
Test Data Class
---------------
Normal     (labe
Abnormal   (labe
---------------
Total: 2911 samp
```

Diagnostic class	Number of subjects
Myocardial infarction	148
Cardiomyopathy/Heart failure	18
Bundle branch block	15
Dysrhythmia	14
Myocardial hypertrophy	7
Valvular heart disease	6
Myocarditis	4
Miscellaneous	4
Healthy controls	52

Hardware Implementation

- Xilinx Spartan 3E + ISE Design Suite
- Design Flow:
  - Quantize model, synthesize, route + place, iterate until things work smoothly.



Key Metric: % speedup in inference on FPGA vs. M1 Neural Engine

Next Steps

- Implement simple NN on FPGA using OSS design tools (e.g. hls4ml)
- Implement more advanced NN on FPGA platform
    - Iterate on CNN model + model quantization + HLS until this works very well.

Team is committed to completing significant work over the 3 weeks of Winter Break

Pierre’s 6 Slides

Hardware Decision

Spartan 3E XA3S500E

- 10,000 logic cells
- MicroBlaze softcore CPU
  - onboard flash memory holds up to 360 kB of C code
- Vivado Design Suite can synthesize a SystemC model into C and VHDL
  - Additional practice in SystemC gained from COE838
- I/O is _____

Dataflow

based on “ECG-based machine-learning algorithms for heartbeat classification”, Nature

Peak Detection → Feature Extraction → Classification

Wavelet Transform

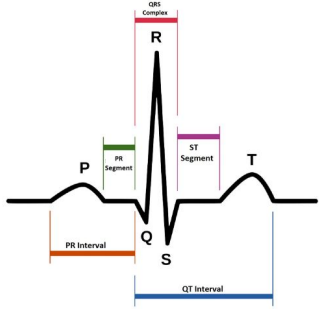
Fractional Fourier Transform

$$F^{\alpha}(x(t)) = X_{\phi}(u) = \int_{-\infty}^{\infty} x(t) K_{\phi}(t,u) dt$$

CNN

1 × 360 data input

$$W_{\psi}(j,k) = \frac{1}{\sqrt{M}} \sum_{k=0}^{M-1} x(t) \psi_{j,k}(t)$$



- Normal
- Ventricular

DFT in VHDL

Ayoub’s 6 Slides

ECG Arrhythmia Detection & My Project Contribution

- ECG arrhythmias occur when the heart beats irregularly (abnormal rhythm).
- Project goal: classify heartbeats as **Normal (N)** vs **Ventricular (V/E)**.
- Ventricular beats are clinically important because they appear wide and abnormal.
- My responsibility: **Develop Python-based CNN and dataset pipeline** for baseline classification.

Dataset Construction (MIT-BIH)

- Used MIT-BIH Arrhythmia Database
- Extracted around-R-peak segments (360 samples)
- 2 classes: Normal (N), Ventricular (V/E)
- Preprocessing: normalization, train–val split
- Output: X_train.npy, y_train.npy, X_val.npy, y_val.npy

CNN Design Options Researched

- Traditional ML (SVM / KNN / Random Forest)
- 2D CNN (treating ECG as an “image-like” input)
- 1D CNN (used)

Marly’s 6 Slides