ARTIFICIAL INTELLIGENCE: METHODS & APPLICATIONS
Self-Driving Car Using Convolutional Neural Networks

Abel Augustine aaugus24@uic.edu
Joseph Procopio procopi2@uic.edu

May 10, 2018

Description

- Cloning Human Behaviour from images
  - Have an AI play a game without any information other than screenshots
- Try to learn, formulate and generalize ideas not explicitly taught.
- Allows us to model real world scenarios.
- A step in the direction of scene understanding.

Motivation

- Scene understanding research. Gathering insights into user behaviour.

- Image Recommendation

- Modelling a world where AI exhibits human characteristics. (Learning human nuances; see Google Duplex)

- Comma.ai

Relevant Methods

- Python Plays GTA V
  - Motivating factor
  - Used AlexNet for training
  - GrayScale Images
  - 3 Classes - AWD

- Improvements:
  - Various screen capture techniques. (We currently capture at above 30fps)
  - Deeper and faster networks
  - Capture in RGB
  - Uses 4/9 class separation and tests with multilabel prediction.
  - Balancing data to lowest of 9 classes:
    - Left, Right, Forward, Brake, Forward Left, Forward Right, Brake Left, Brake Right, No Op
  - Merge data into less files for faster training

Technologies
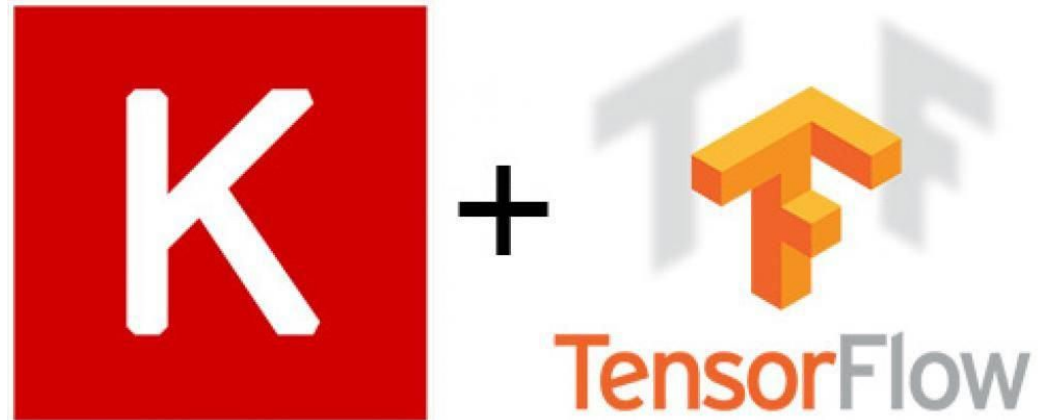
Data - A Love/Hate Relationship

- Captured from us playing the game
  - Screenshots ~30 FPS
  - Keystrokes for BR, FW, L, R etc. mapped to screenshot
  - RGB Images
  - Stored in a numpy array and saved every 1,000 frames
  - About 50 seconds of gameplay is 300 MB!

- Balancing Data,
  - Unbalanced data heavily biased to forward only
  - Balance to lowest of 9 and 4 classes.
  - Merge files from 100+ to 8
  - Files drop from 300 MB to 100 MB and as low as 50 MB
  - Trained over 45-120 minutes of gameplay across different models.

Data Collected from :
Mayur (Main, fast driving), Joseph, Mayur (Main, accurate driving), Abel

Sample Solution

Cloning human behaviour! (or how to play F1)

Resize

320x240

Save with class labels to train

Sends keep press to the output stream.

AlexNet/ MobileNet

Handling the challenge of class imbalance.

Relevant research

MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications (MobileNet)

ImageNet Classification with Deep Convolutional Neural Networks (AlexNet)

Going Deeper with Convolutions (Inception models)

Rethinking the Inception Architecture for Computer Vision (Improved versions of Inception)

DeepDriving: Learning Affordance for Direct Perception in Autonomous Driving
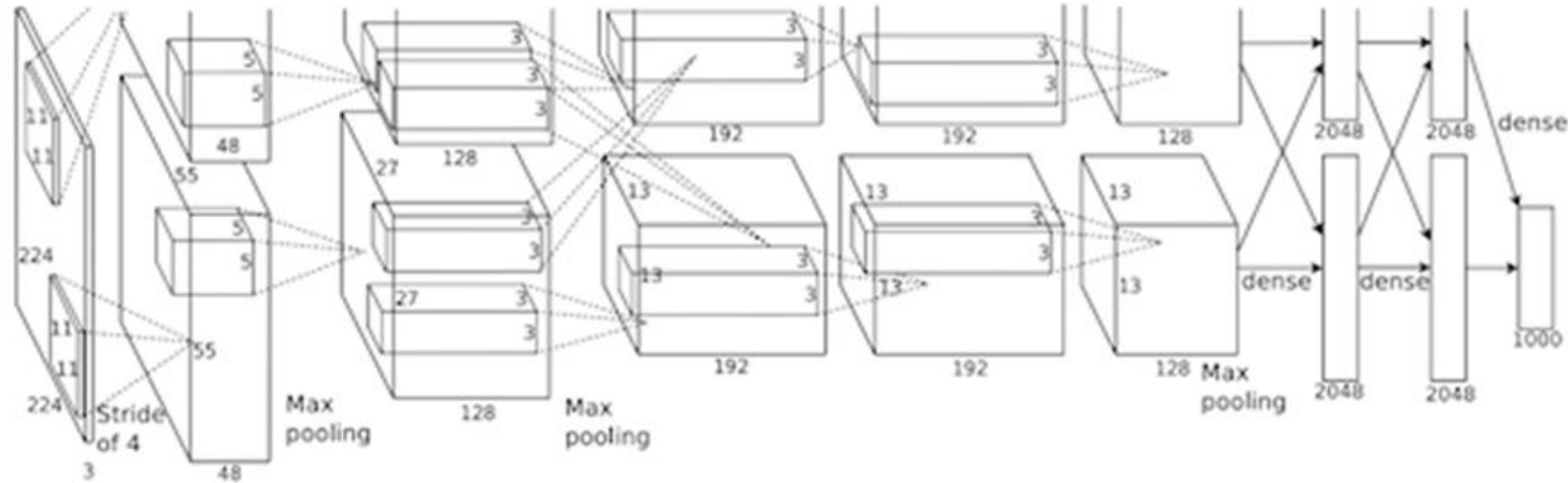
End to End Learning for Self-Driving Cars

Alvinn: An autonomous land vehicle in a neural network

Neural network perception for mobilerobot guidance

Architecture-1 Alexnet

AlexNet architecture (May look weird because there are two different "streams". This is because the training process was so computationally expensive that they had to split the training onto 2 GPUs)
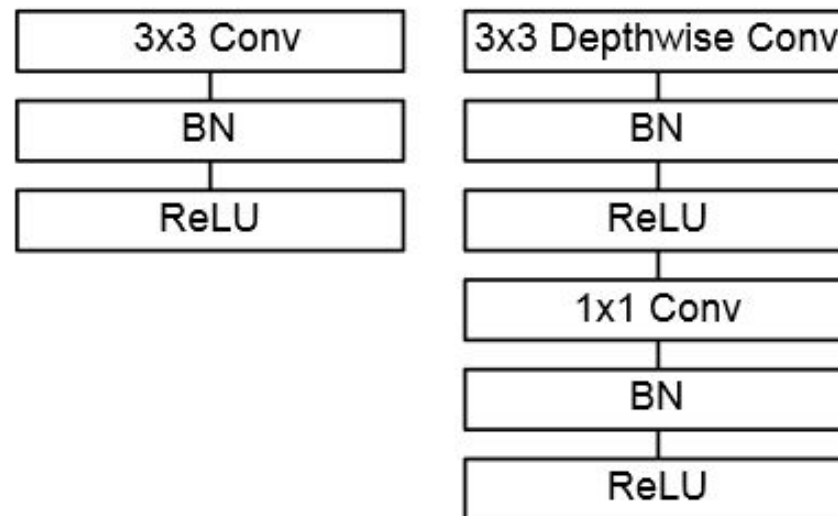
Architecture-2 MobileNet

Figure 3. Left: Standard convolutional layer with batchnorm and ReLU. Right: Depthwise Separable convolutions with Depthwise and Pointwise layers followed by batchnorm and ReLU.

Models Trained

Data: 47 minutes of gameplay (pre-balancing)

- Alexnet (75 epochs, 4 classes, old balancing)
- MobileNet (~75 epochs, 4 classes, old balancing)

With Increased Data (~3x)

- MobileNet (~30 epochs, 9 classes, new balancing)
- MobileNet (~30 epochs, 4 classes, new balancing)

Metrics

- Validation and test set: Categorical Accuracy per batch file.

- Average accuracy per epoch.

- Cross-validation accuracy. (Accuracy = Categorical Accuracy)
    - We also inspected Precision, Recall, F1 score to correct imbalance classes.
    - Hyperparameters - the resolution of images required, output classes, RGB vs Grayscale, balancing threshold.

Future improvements

- Try other convolutional networks like Inception series etc.

- Run with more GPU memory, RAM etc.

- Larger dataset for multiclass single label.

- Custom loss function to handle imbalance.

- Combine other features to provide context to network.

- Data Creation: Playing for Data: Ground Truth from Computer Games
  https://arxiv.org/pdf/1608.02192.pdf or GAN

Code Review & Demo

- Go through entire code base.
- Predict while playing F1 2017