Distracted Driver Detection

CS 747 Semester Project

Angeela Acharya
Anita Tadakamalla
Sulabh Shrestha

Problem Statement

- 1 out of 5 accidents due to distracted driving
- Initiative by State Farm:
  - Use of Dashboard camera for automatic detection of distracting behaviors
  - Classification of each driver's behavior
- "State Farm Distracted Driver Detection"
  - Kaggle competition to predict the driver's action in each image
  - Dataset: 10 possible actions (classes) - driving safe, texting, talking, drinking, etc.

(a) C0: Drive safe (b) C1: Text left (c) C2: Talk left (d) C3: Text right (e) C4: Talk right (f) C5: Adjust radio (g) C6: Drink (h) C7: Hair and makeup (i) C8: Reaching behind (j) C9: Talk to passanger

Approach

We experimented with the following:

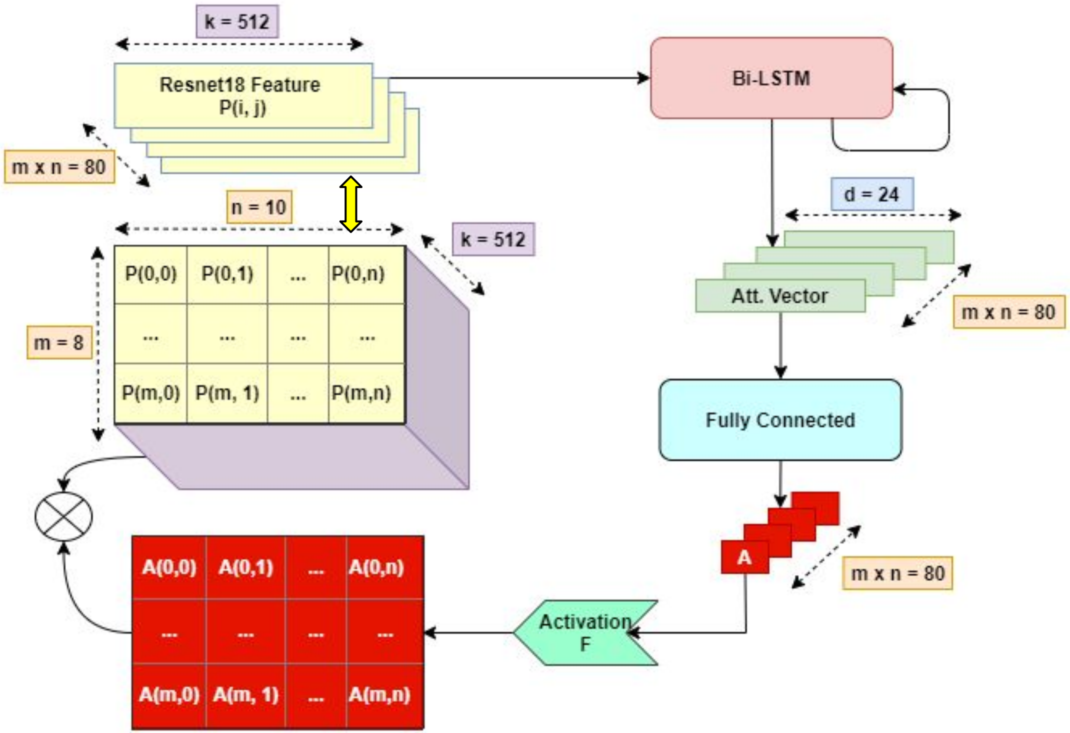
- Building classification models
    - Compare performances of deep learning classification technique
- Attention
    - Focus on relevant spatial location
    - Use importance weighted features
    - Offer some interpretability
- Knowledge Distillation
    - “Deeper the Better” holds only on high capacity GPUs
    - Only smaller models feasible in cars
    - Transfer generalization ability from bigger models to smaller models

Approach: Classification Models

- We experimented with the following classification models:
  - VGG-16, VGG-19, InceptionV3, Xception, ResNet-18, ResNet-101
- Evaluation Metric:
  - Accuracy
  - Logarithmic loss
- Issues: Overfitting
- Solutions:
  - Image Augmentation
  - Ensemble model -- Mean Ensembling: posterior probability is calculated as the mean of the predicted probabilities from the component models

Approach: Attention Networks

- Weight features based on relevance of location
- ResNet 18 Features (m x n x 512)
- P(i,j)
  - 1 x 1 x 512 = 1 time-step
  - m x n locations
- Bi-LSTM
  - Look at all locations P(i,j)
  - Produce appropriate representation
- Fully-Connected
  - Use Bi-LSTM representation
  - Generate 1 attention weight A(i, j)
  - Activation (F):
    - ReLU + L1 Norm
    - Softmax
- Entropy loss for Attention

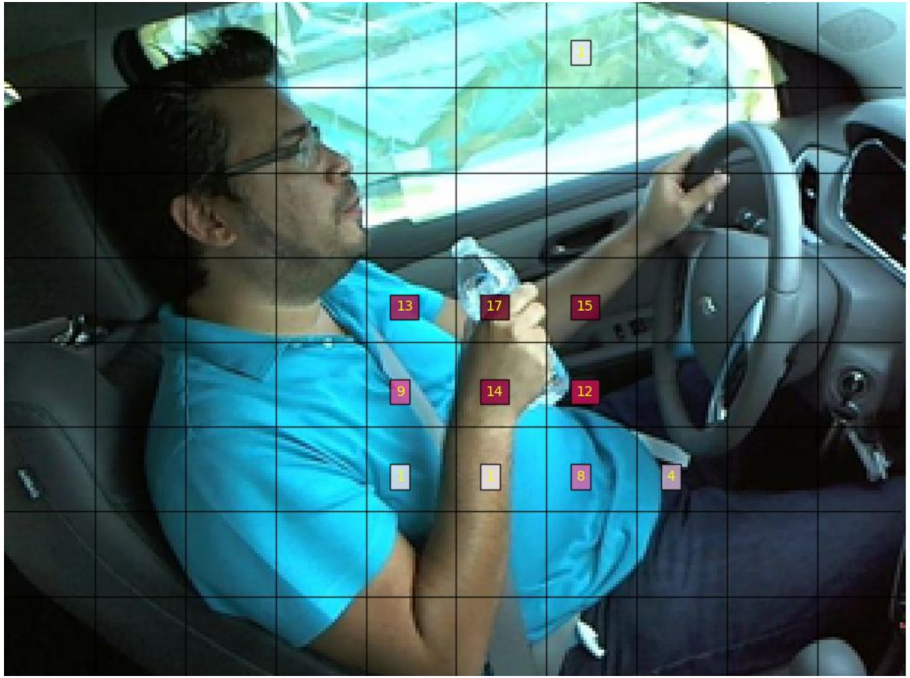


Attention Visualization (Correct Results)

➔ Round(Attention, 2)*100 > 0

➔ Model: F=Softmax + Conv. + Entropy

Label: Text Right
Attention: around <u>Phone</u>

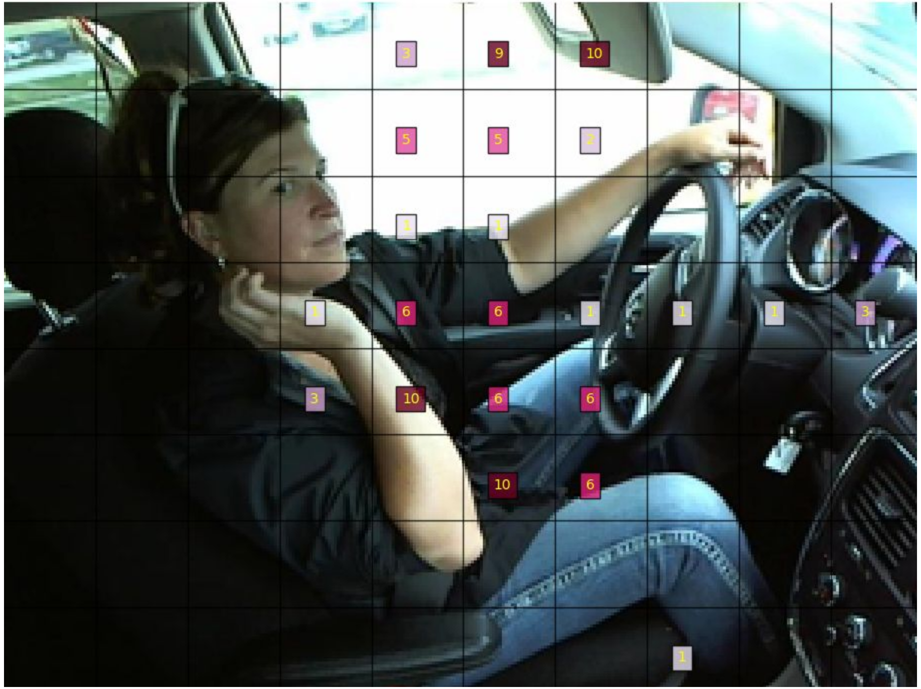


Label: Drinking
Attention: around <u>Bottle</u>

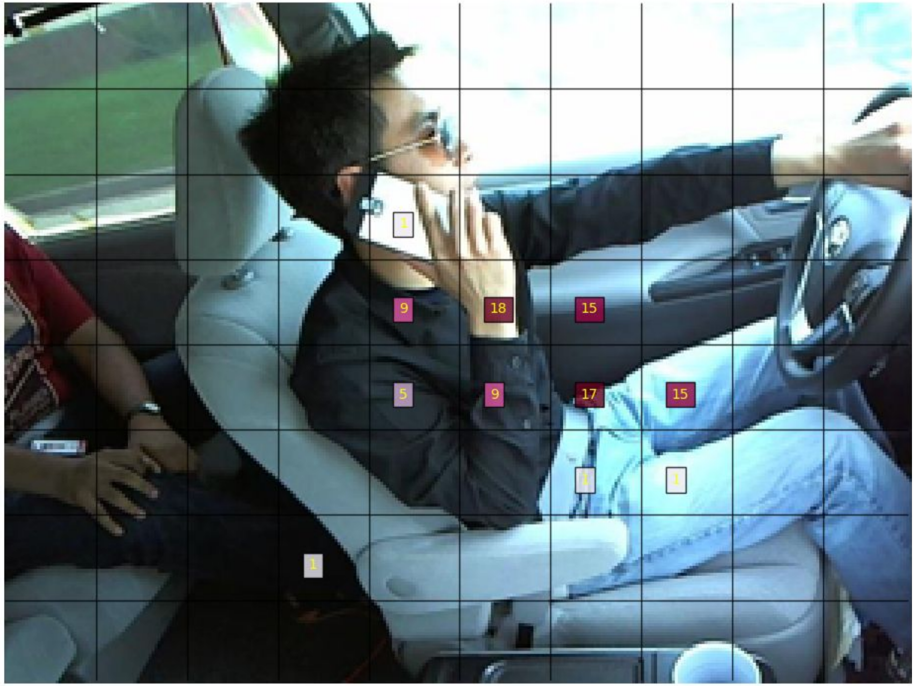
Attention Visualization (Correct Results)

➔ Round(Attention, 2)*100 > 0

➔ Model: F=Softmax + Conv. + Entropy



Label: Hair and Makeup
Attention: around Mirror and Hand

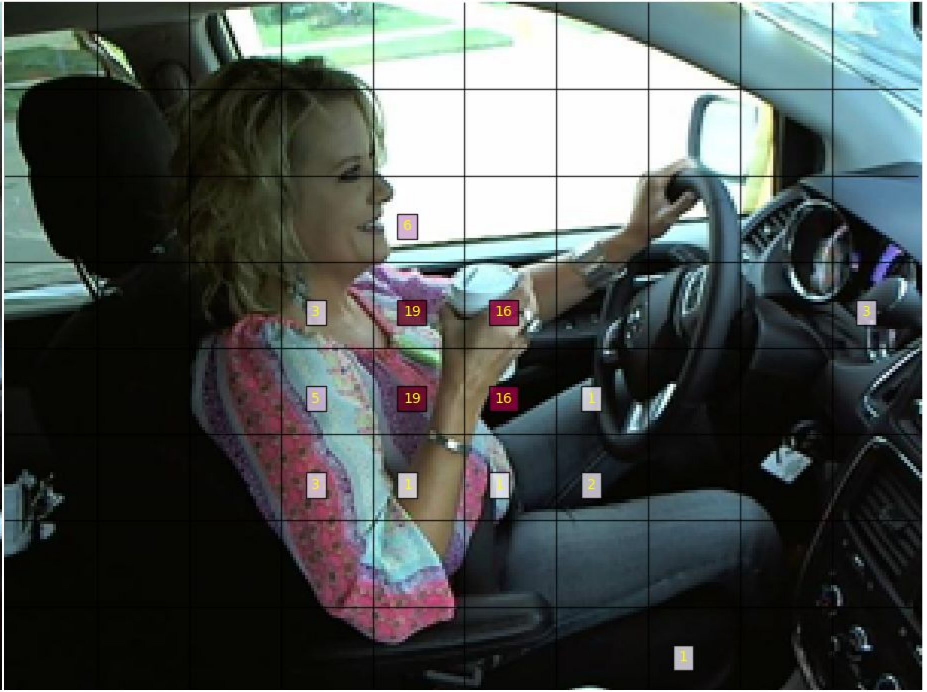


Label: Talking on Phone - Right
Attention: around Phone (Big Receptive Field)

Attention Visualization (Incorrect Results)

- ➔ Round(Attention, 2)*100 > 0
- ➔ Model: F=Softmax + Conv. + Entropy

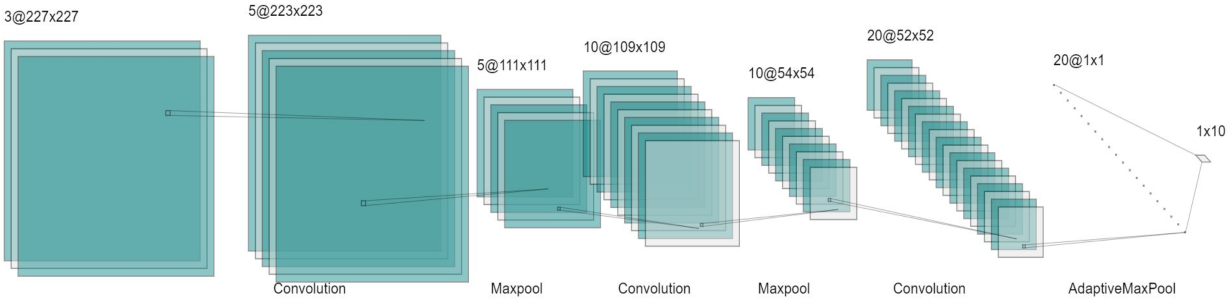
Label: Reaching Behind
Attention: Wrong Location



Label: Talking on Phone - Right
Attention: Right location but Wrong label

Approach: Knowledge distillation

- Teacher networks: ResNet-18 and ResNet-101
- Student network:



Our approach:

- Train teacher network
    - Try different temperatures, T = {1, 4, 8, 12}
    - Increased Entropy → Softer Probabilities
- Optimize Student network
    - Soft cross entropy loss: Soft probabilities from teacher network as labels
    - Hard cross entropy loss: One-Hot targets

$$q_i = \frac{exp(z_i/T)}{\sum_j exp(z_j/T)}$$

Quantitative Results

Metric	Network					Attention				Knowledge Distillation Students (T=4)		
	VGG-16	VGG-19	Xception	ResN 18	ResN 101	Soft+ Conv	Soft+ Conv+ Ent.	ReLU+ Conv	Soft+ Linear	ResN 18	ResN 101	Plain
Train loss/ Entropy	0.41	0.97	0.47	0.17	**0.001**	3.7304	**2.419**	3.789	3.321	**0.17**	0.65	0.36
Val loss	0.56	0.82	0.55	0.16	**0.01**	0.044	0.028	**0.020**	0.031	**0.30**	0.99	0.4
Val acc	0.83	0.76	0.83	0.99	**0.996**	0.995	0.996	0.996	**0.997**	**0.90**	0.84	0.86
Test loss	-	-	-	0.49	**0.38**	0.829	**0.670**	0.804	1.000	1.5	2.3	**1.16**

- Network: ResNet-101 best score in kaggle (Top 16%)
- Distillation: ResNet18 better than ResNet-101 as teacher
- Attention: Softmax Activation with Entropy loss gives best performance

Contributions

- Angeela: Implementation of Knowledge distillation technique, training and validation of ResNet-101 and ResNet-18 models
- Anita: Training and validation of various models, attempted implementation of ensemble model
- Sulabh: Architecture design, implementation and training of the Attention Networks

Thank You