



---

# Computer Vision of




## Distracted Driver Detection



Group 3: Zhuohan Yu, Zeqiu Zhang

---



# Background

Distracted driving is one of the main reasons for car accidents.

According to CDC's data, it causes about 425,000 people injured and 3,000 people killed every year.



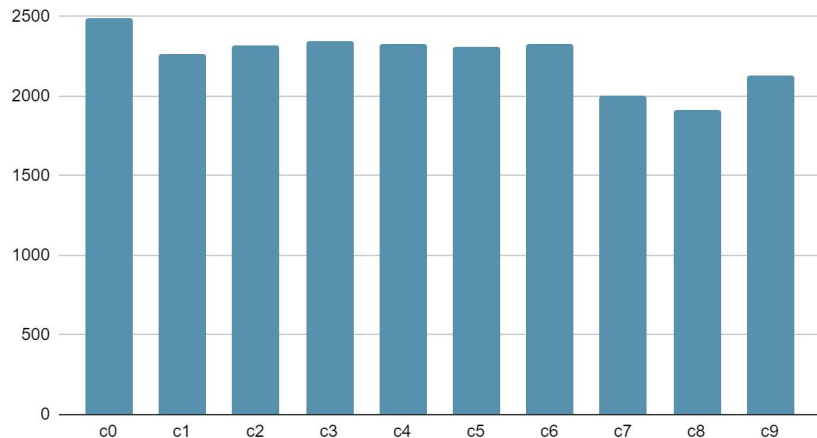
# Dataset Description

Train: 19060

Test: 79726

Pixel Size: 460 \* 680

Train Data Distribution



## ***10 Classes:***

c0: safe driving

c1: texting - right

c2: talking on the phone - right

c3: texting - left

c4: talking on the phone - left

c5: operating the radio

c6: drinking

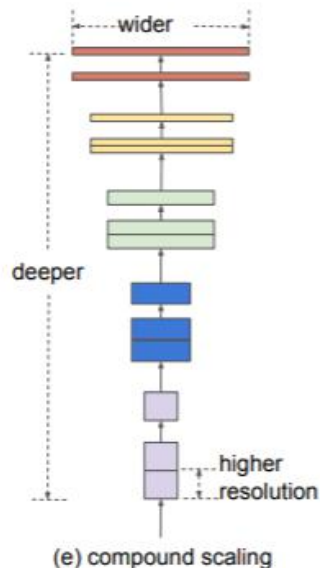
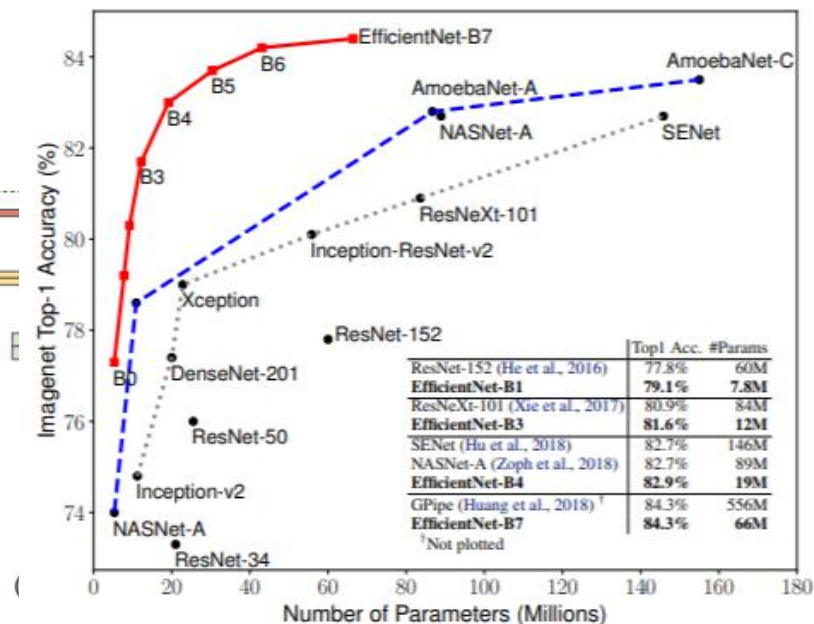
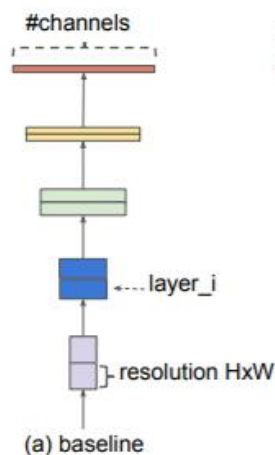
c7: reaching behind

c8: hair and makeup

c9: talking to passenger

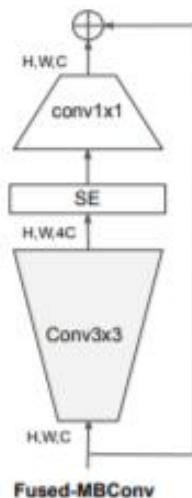
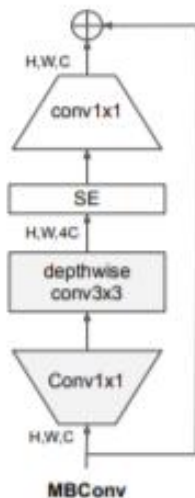
# Model Description

## EfficientNet



# Model Description

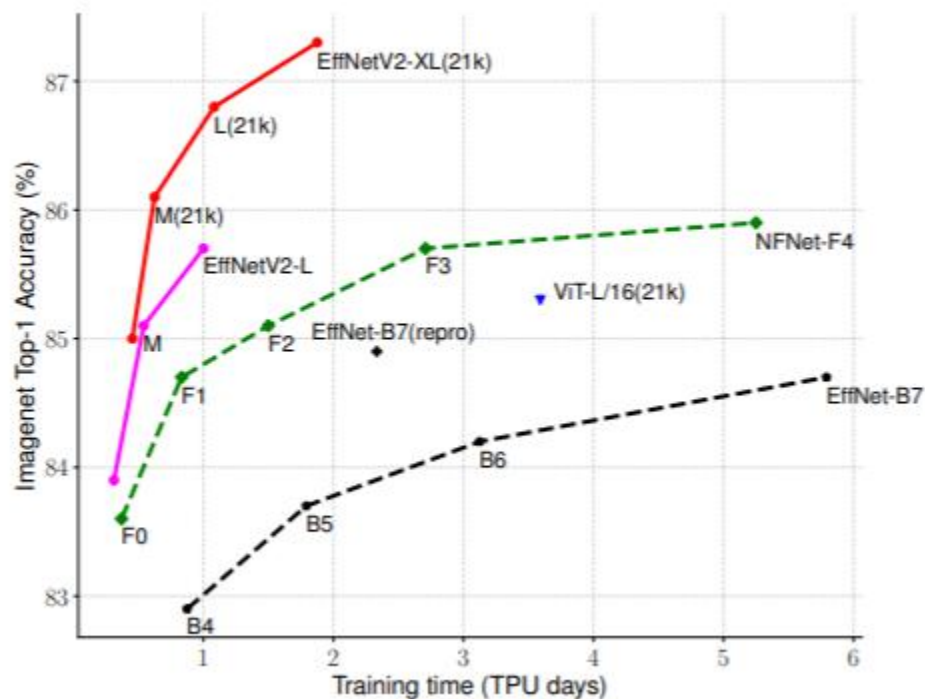
## EfficientNetV2



Stage	Operator	Stride	#Channels	#Layers
0	Conv3x3	2	24	1
1	Fused-MBConv1, k3x3	1	24	2
2	Fused-MBConv4, k3x3	2	48	4
3	Fused-MBConv4, k3x3	2	64	4
4	MBConv4, k3x3, SE0.25	2	128	6
5	MBConv6, k3x3, SE0.25	1	160	9
6	MBConv6, k3x3, SE0.25	2	256	15
7	Conv1x1 & Pooling & FC	-	1280	1

# Model Description

## EfficientNetV2



# Experiment Setup

1. Batch challenge
2. Two ways of split
  - a. Regular Split
  - b. Kfold Split
3. Model definition
  - a. Pretrained Models
  - b. Optimizers
  - c. Loss function
  - d. Callbacks
    - i. checkpoint
    - ii. early stopping
    - iii. ReduceLROnPlateau

Framework:  
*Tensorflow*

# Experiment Setup

1. Tuning parameters
  - a. dropout rate
  - b. number of epoch
  - c. image size
2. Double Ensemble
  - a. Kfold Ensemble
  - b. Model Ensemble



# Results

1. Resnet50, Resnet152V2
2. Densenet
3. InceptionV3
4. EfficientNetV2B2
5. EfficientNetV2B3

Pretrained Model	Val_accuracy	Private Score
Densenet121	0.9969	0.32719
	0.9967	0.36186
Densenet169	0.9971	0.34589
Densenet201	0.9967	0.36915
InceptionV3	0.9946	0.28210
	0.9958	0.37053
	0.9943	0.39464
Resnet152V2	0.9962	0.59611
EfficientNetB2	0.9969	0.34956
EfficientNetV2B2	0.9967	0.31996
	0.9969	0.29559
	0.9962	0.35109
	Kfold Ensemble	0.20779
EfficientNetV2B3	0.9965	0.30356
	0.9966	0.30581
	Kfold Ensemble	0.22650
EfficientNetV2M	0.9958	0.32306
EfficientNetV2L	0.9944	0.38809
<b>Model Ensemble</b>		<b>0.20481</b>

# Conclusions

1. EfficientNetV2B2 and EfficientNetV2B3 perform the best among all the models
1. Kfold and ensemble could improve the score

# Future Researches

1. Use more sophisticated augmentation like CutMix or Mixup
2. Try larger image size for EfficientNetV2M and EfficientNetV2L
3. Feed information of timeline from same drivers' pictures into model

# Reference

<https://www.kaggle.com/c/state-farm-distracted-driver-detection/data> Kaggle: State Farm Distracted Driver Detection

Tan, M., & Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning* (pp. 6105-6114). PMLR.

Tan, M., & Le, Q. V. (2021). Efficientnetv2: Smaller models and faster training. *arXiv preprint arXiv:2104.00298*.

*State Farm distracted driver detection*. Kaggle. (n.d.). Retrieved December 6, 2021, from <https://www.kaggle.com/c/state-farm-distracted-driver-detection/discussion/22631>.

*State Farm distracted driver detection*. Kaggle. (n.d.). Retrieved December 6, 2021, from <https://www.kaggle.com/c/state-farm-distracted-driver-detection/discussion/22906>.

*State Farm distracted driver detection*. Kaggle. (n.d.). Retrieved December 6, 2021, from <https://www.kaggle.com/jiaodong/vgg-16-pretrained-loss-0-23800>.

**Thank you**