

Comparing CNN Architectures for Visual Impairment Assistive Technology Applications

Allen Shelton

Outline

1. Problem Statement
2. Introduction
3. Description of
Solution/Implementation
4. Results

Problem Statement

- 295 million people globally have moderate to severe visual impairment, 43.3 million people legally blind
- Reduced quality of life
- Deep learning solutions for understanding of environment



Introduction

- Improvement on results of previous work
- Visual understanding in context of college campus using webcam and text-to-speech
- Used transfer learning with AlexNet
- 3 methods to try to improve performance:
 - Collect more data
 - Different CNN architectures
 - Data Augmentation

AlexNet changes from original work

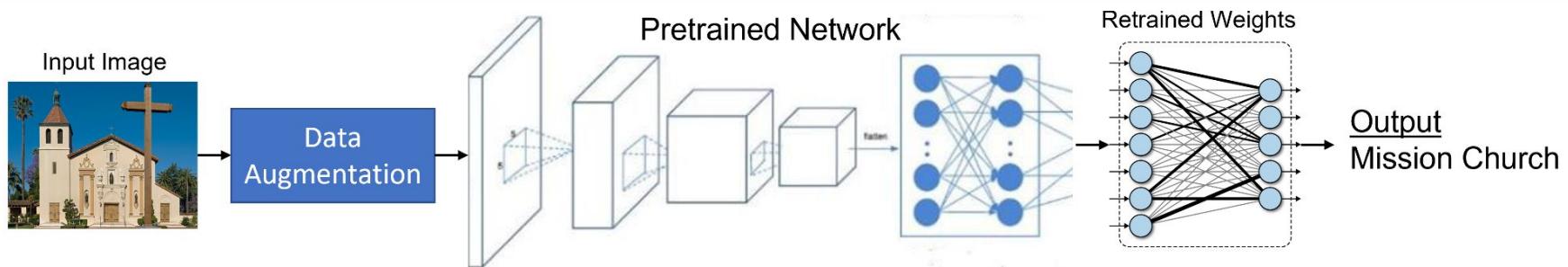
Original AlexNet		Modified AlexNet	
Layer Type	Layer Details	Layer Type	Layer Details
Fully Connected	4096 fully connected layer	Fully Connected	4096 fully connected layer
ReLU	ReLU	ReLU	ReLU
Dropout	50% dropout	Dropout	50% dropout
Fully Connected	1000 fully connected layer	Fully Connected	2048 fully connected layer
Softmax	softmax	ReLU	ReLU
Classification Output	crossentropyex	Dropout	50% dropout
		Fully Connected	10 fully connected layer
		Softmax	softmax
		Classification Output	crossentropyex



Description of Solution/Implementation



Block Diagram



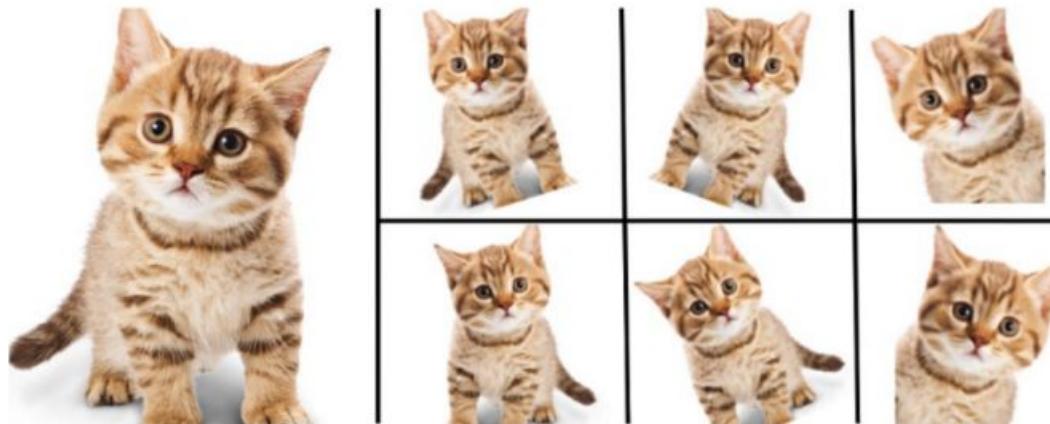
The Dataset

- Group 1: 171 images
- Group 2: 283 images
- More, better quality data = better performance

Group 1	Group 2
Bed	Benson Center
Bench	Bronco Statue
Car	Casa Italiana RLC
Couch	Daly Science
Door	Dowd Art Building
Elevator	Graham RLC
Exit Sign	Kenna Hall
Fire Extinguisher	University Library
Folder/Binder	Mission Church
Fork	O'Connor Hall
Knife	
Palm Tree	
Pen/Pencil	
Stairs	

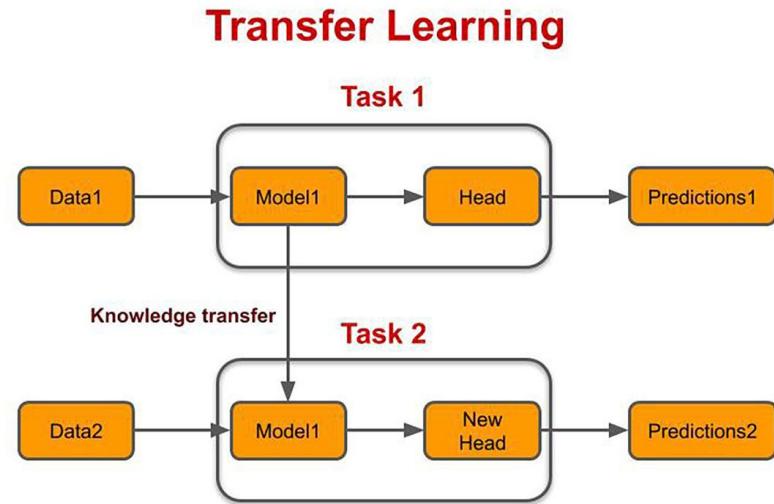
Data Augmentation

- Used in applications with limited data
- Essentially creates more data while preventing overfitting

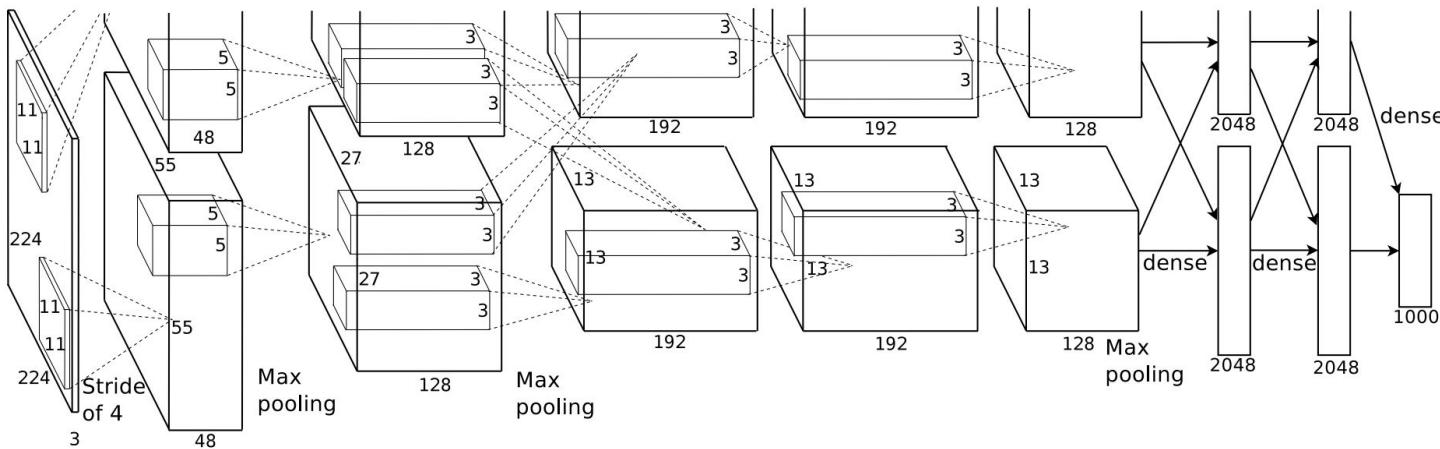


Transfer Learning

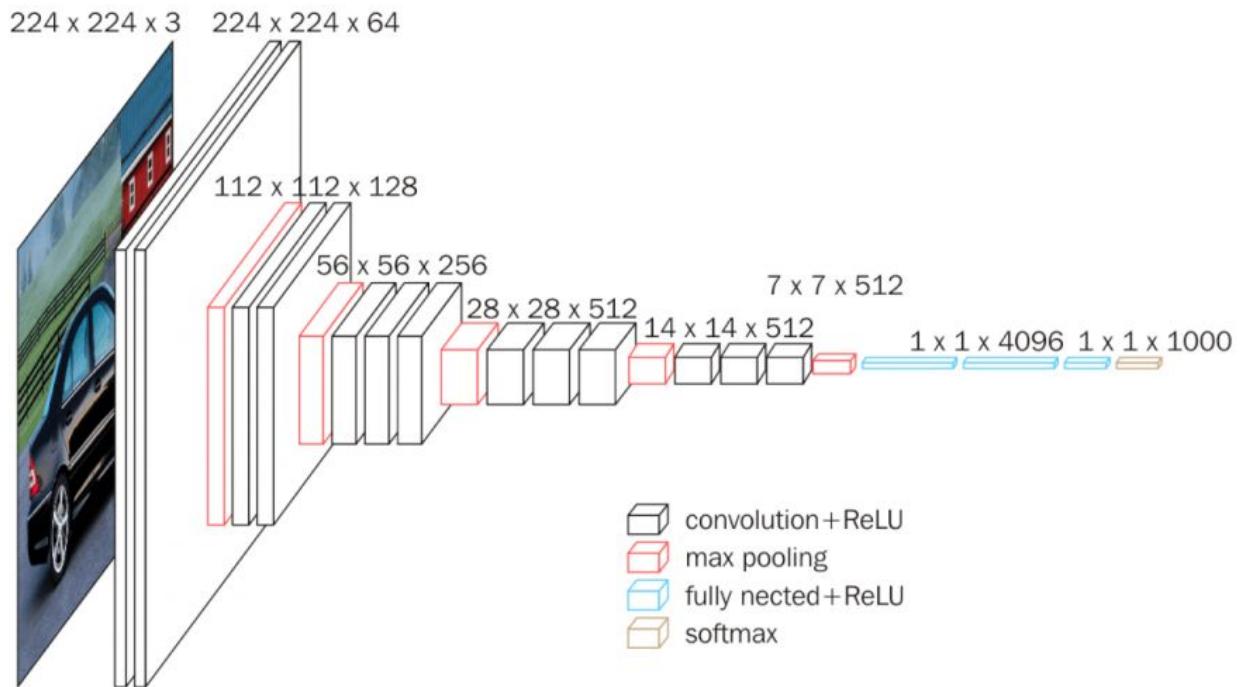
- Designing and training neural networks from scratch can be time-consuming and lead to suboptimal results
- Use pretrained neural network architectures and retrain them for a similar but new task.
- Final layers need to be modified to fit dataset



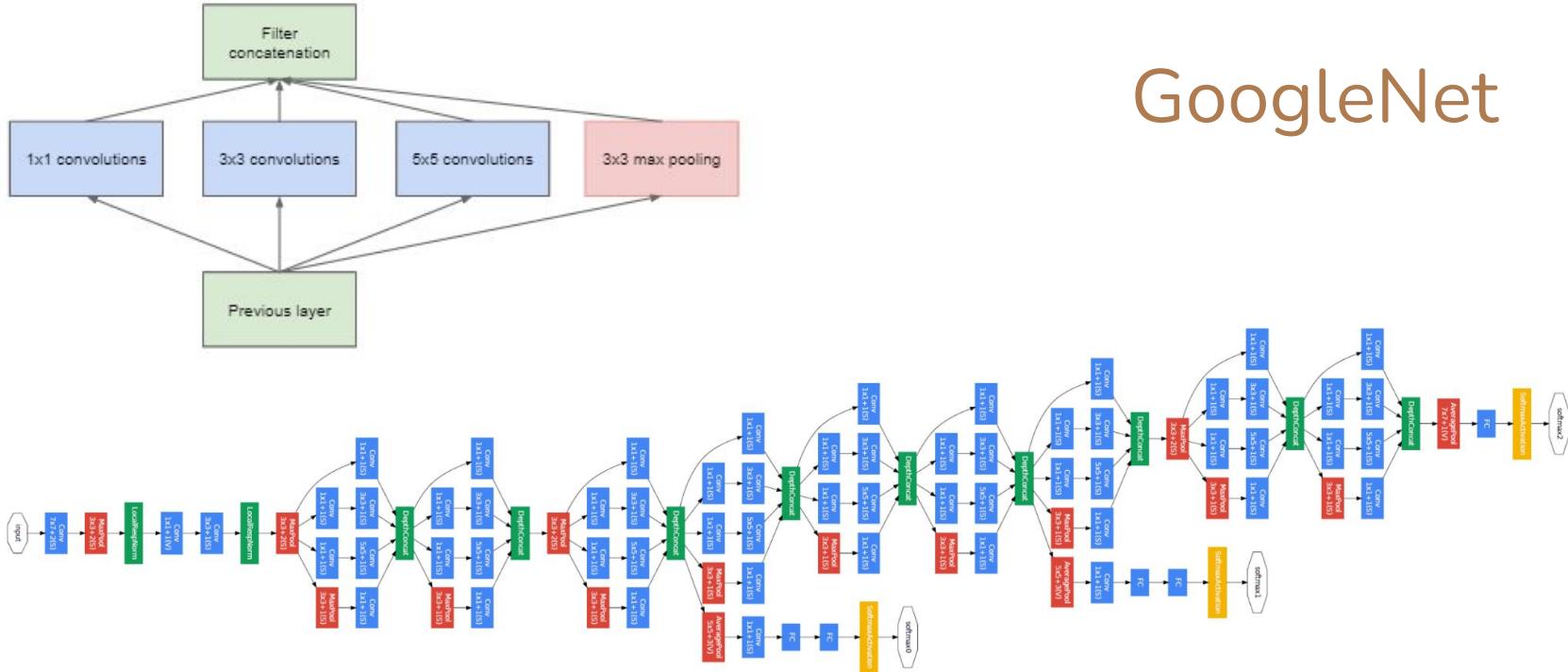
AlexNet



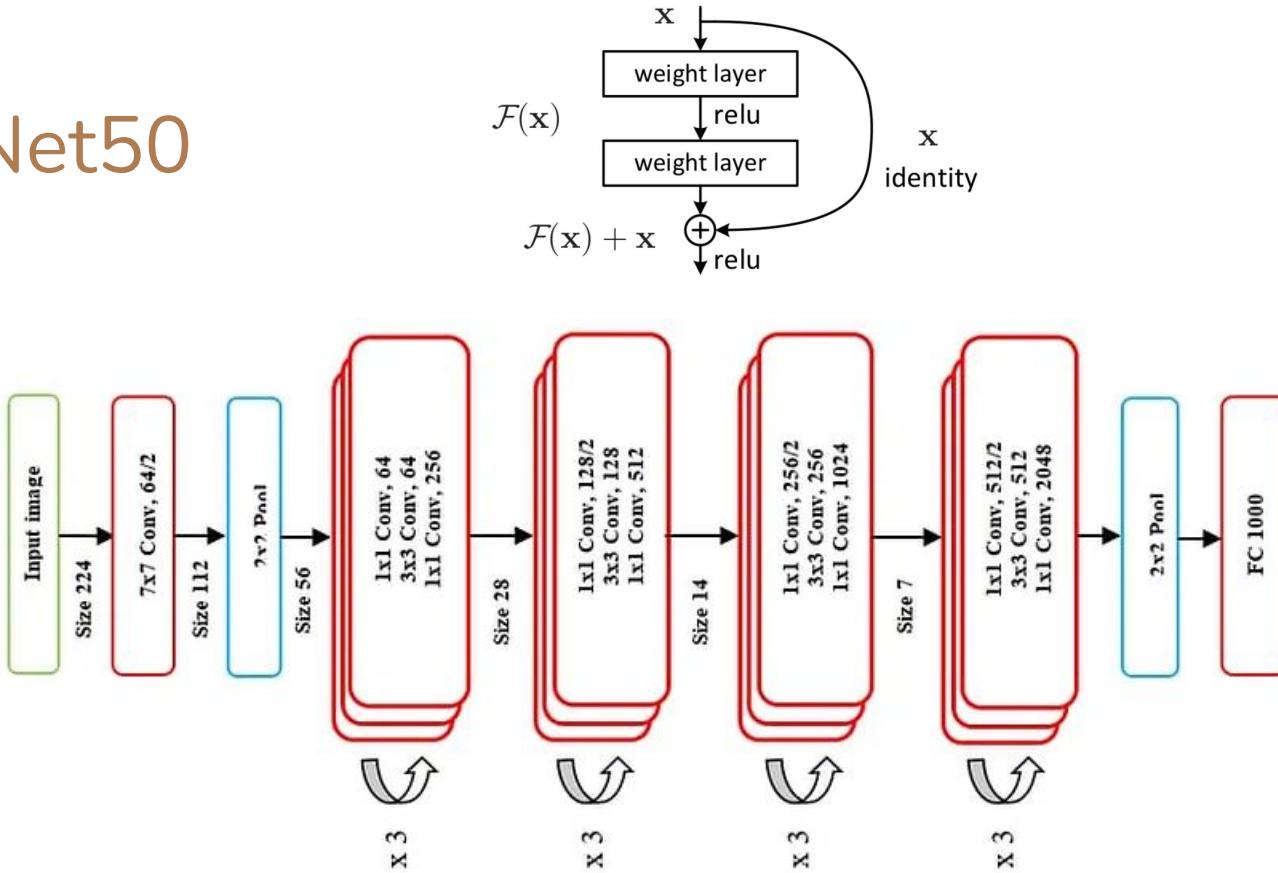
VGG-16



GoogleNet



ResNet50



Results

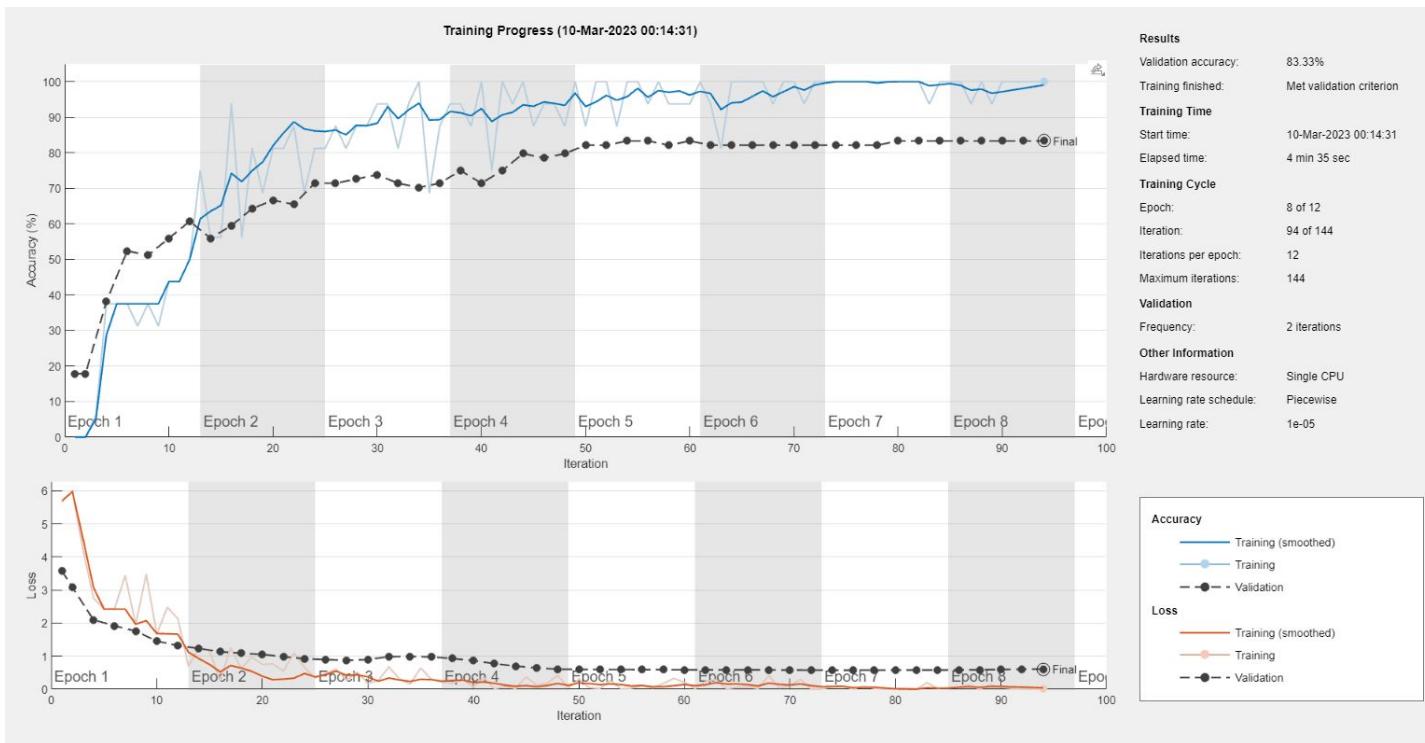
Results for Each Test Case

1. Validation Accuracy/Loss
2. Confusion Matrix
3. Recall, Precision and F1 score for each class
4. ROC Curve for each class
5. Random Selection of Images with predictions

Baseline Results

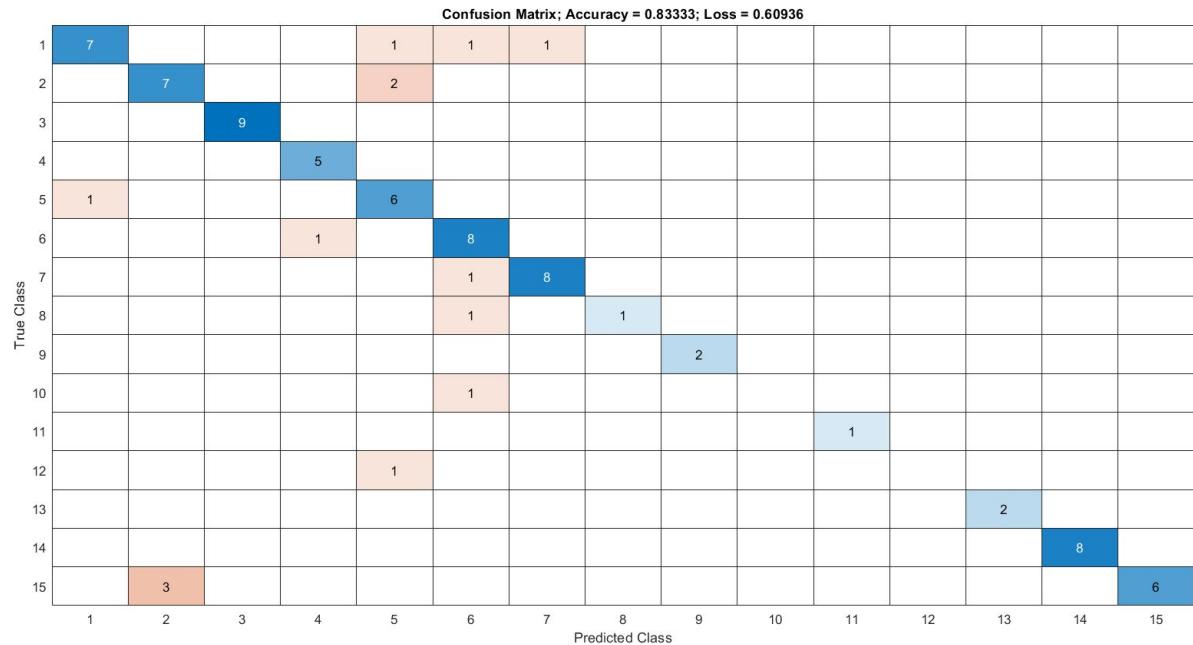
- CNN: AlexNet
- Original Dataset
- Image Augmentation: Random Rotation (-20,20) degrees, Random X-Y Translation 3 pixels
- Initial Learn Rate: 0.0001
- Batch Size: 16
- Input image size: 224 x 224

Gr. 1 Baseline Results: Val. Accuracy/Loss



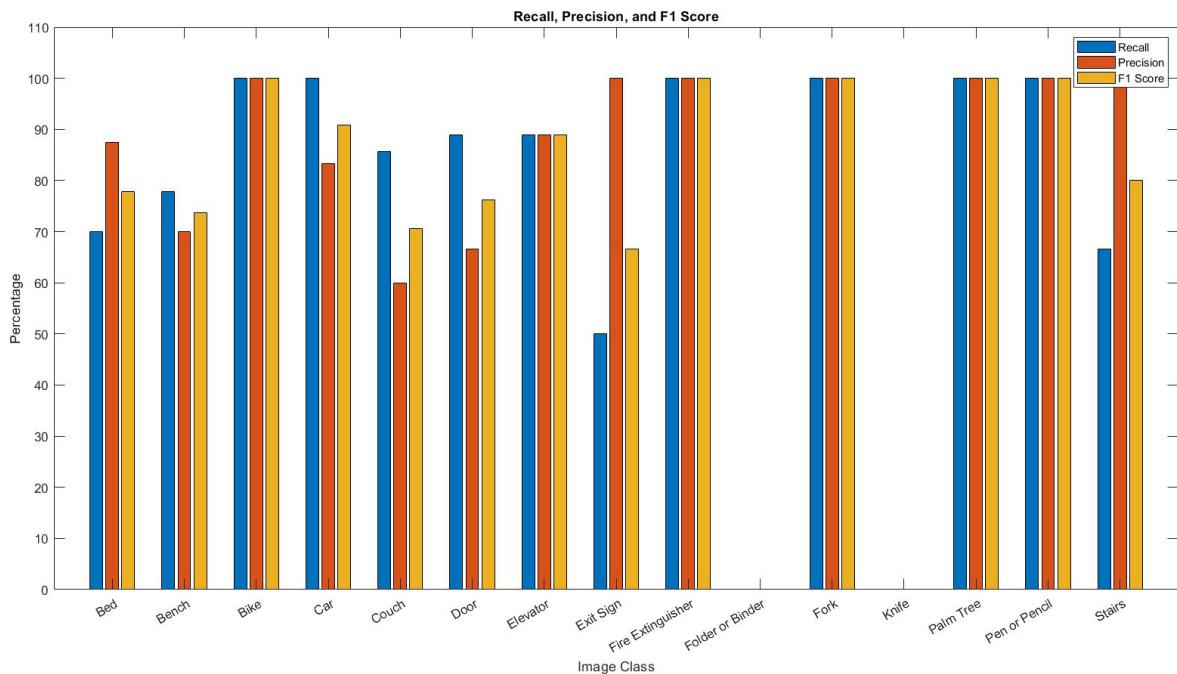
Baseline Results: Confusion Matrix

- | | | | |
|----|--------------|-----|------------|
| 1. | Bed | 10. | Folder/ |
| 2. | Bench | | Binder |
| 3. | Bike | 11. | Fork |
| 4. | Car | 12. | Knife |
| 5. | Couch | 13. | Palm Tree |
| 6. | Door | 14. | Pen/Pencil |
| 7. | Elevator | 15. | Stairs |
| 8. | Exit Sign | | |
| 9. | Fire | | |
| | Extinguisher | | |

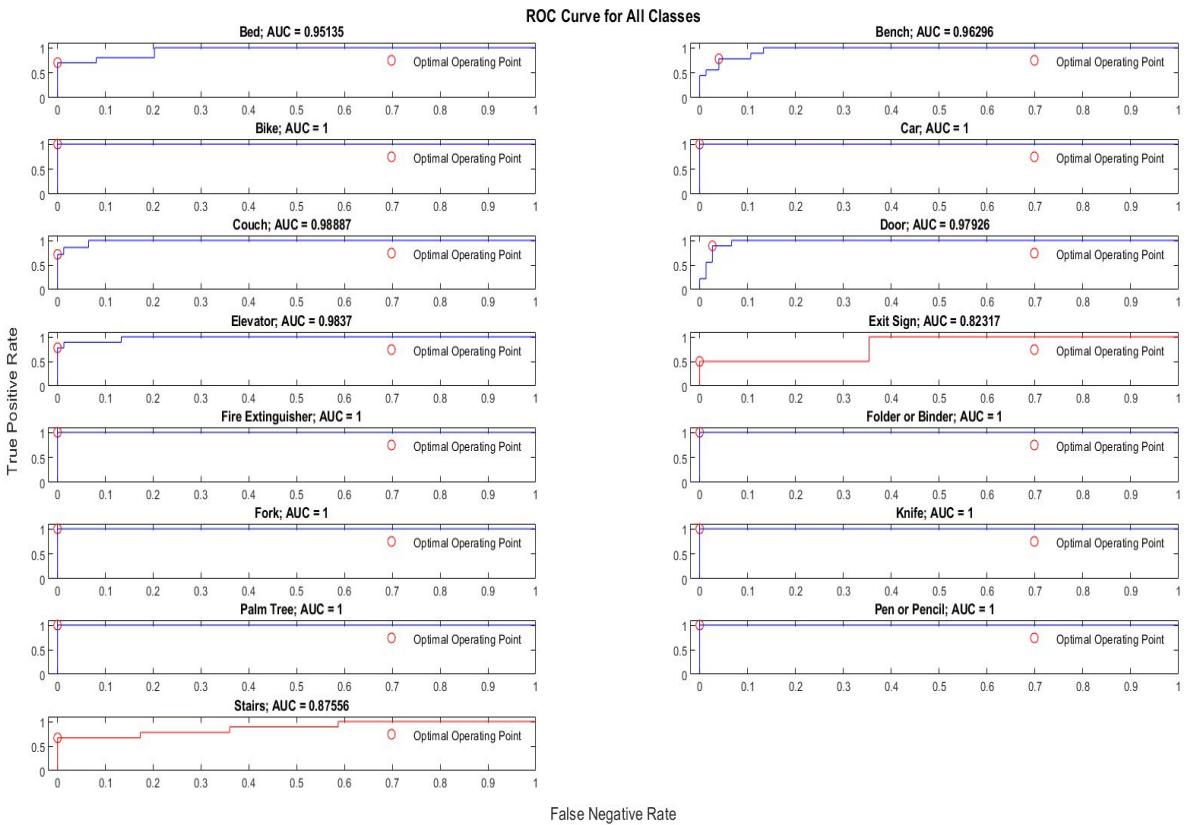


Baseline Results: Recall, Precision, F1 Score

	Recall	Precision	F1
Bed	70	87.5000	77.7778
Bench	77.7778	70	73.6842
Bike	100	100	100
Car	100	83.3333	90.9091
Couch	85.7143	60	70.5882
Door	88.8889	66.6667	76.1905
Elevator	88.8889	88.8889	88.8889
Exit Sign	50	100	66.6667
Fire Extinguisher	100	100	100
Folder or Binder	0	NaN	NaN
Fork	100	100	100
Knife	0	NaN	NaN
Palm Tree	100	100	100
Pen or Pencil	100	100	100
Stairs	66.6667	100	80



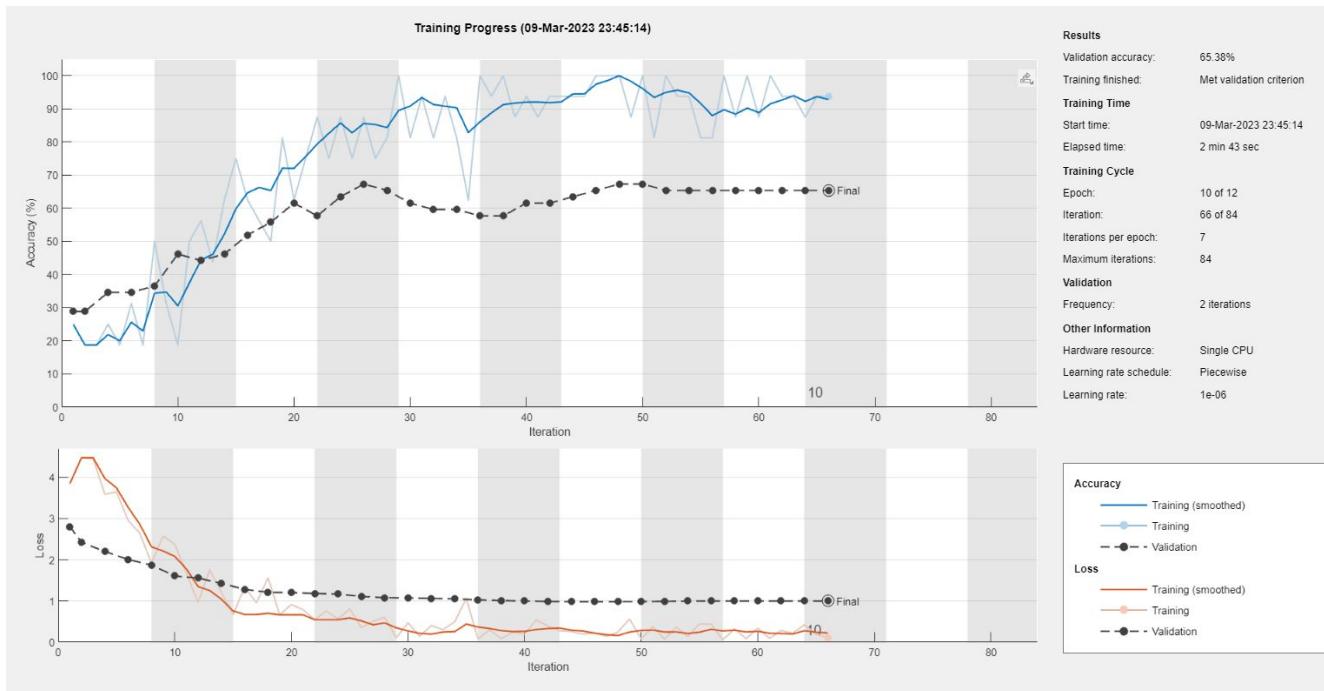
Baseline Results: ROC Curves



Baseline Results: Predictions

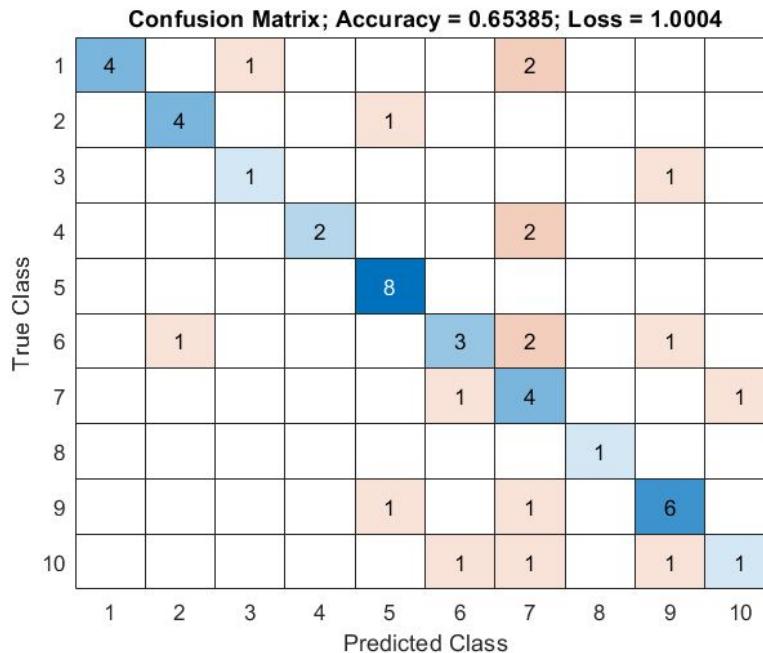


Gr. 2 Baseline Results: Val. Accuracy/Loss



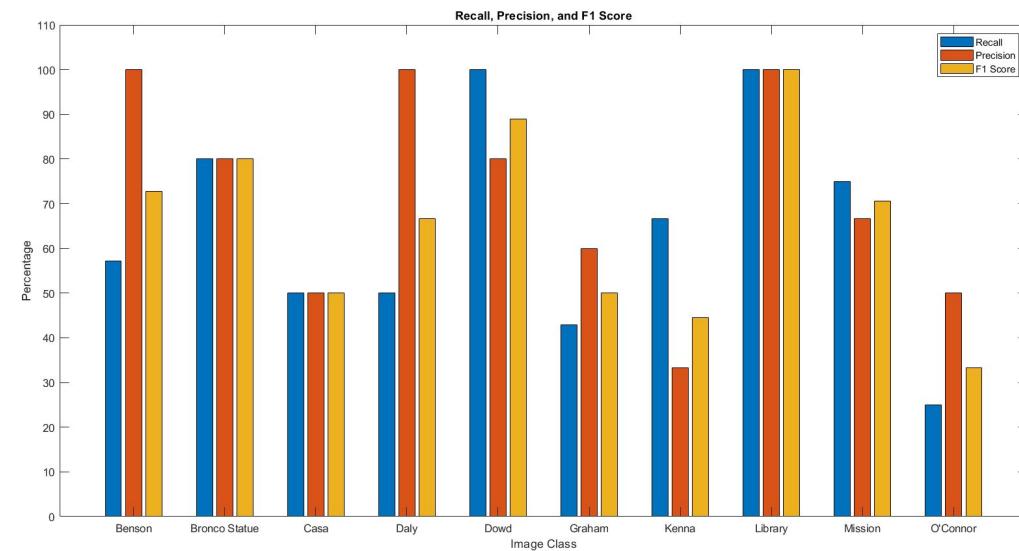
Baseline Results: Confusion Matrix

1. Benson
2. Bronco Statue
3. Casa RLC
4. Daly Science
5. Dowd
6. Graham RLC
7. Kenna
8. Library
9. Mission Church
10. O'Connor

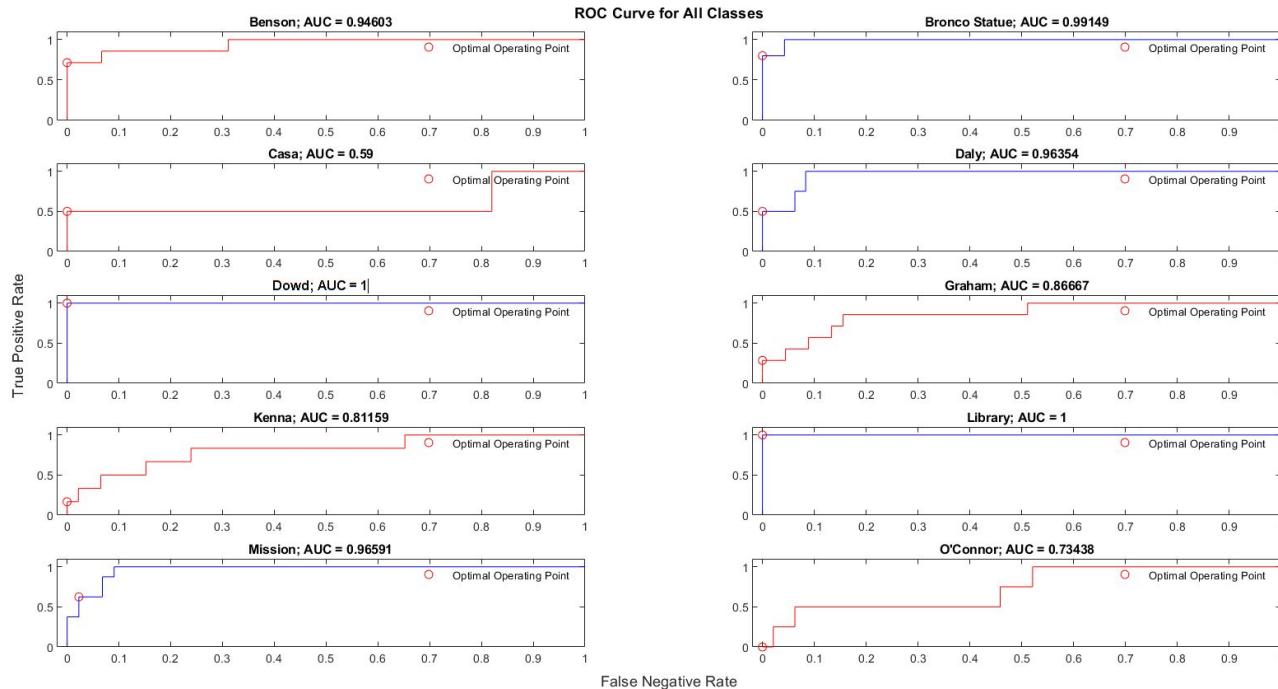


Baseline Results: Recall, Precision, F1 Score

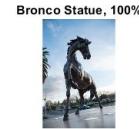
	Recall	Precision	F1
Benson	57.1429	100	72.7273
Bronco Statue	80	80	80
Casa	50	50	50
Daly	50	100	66.6667
Dowd	100	80	88.8889
Graham	42.8571	60	50.0000
Kenna	66.6667	33.3333	44.4444
Library	100	100	100
Mission	75	66.6667	70.5882
O'Connor	25	50	33.3333



Baseline Results: ROC Curves



Baseline Results: Predictions



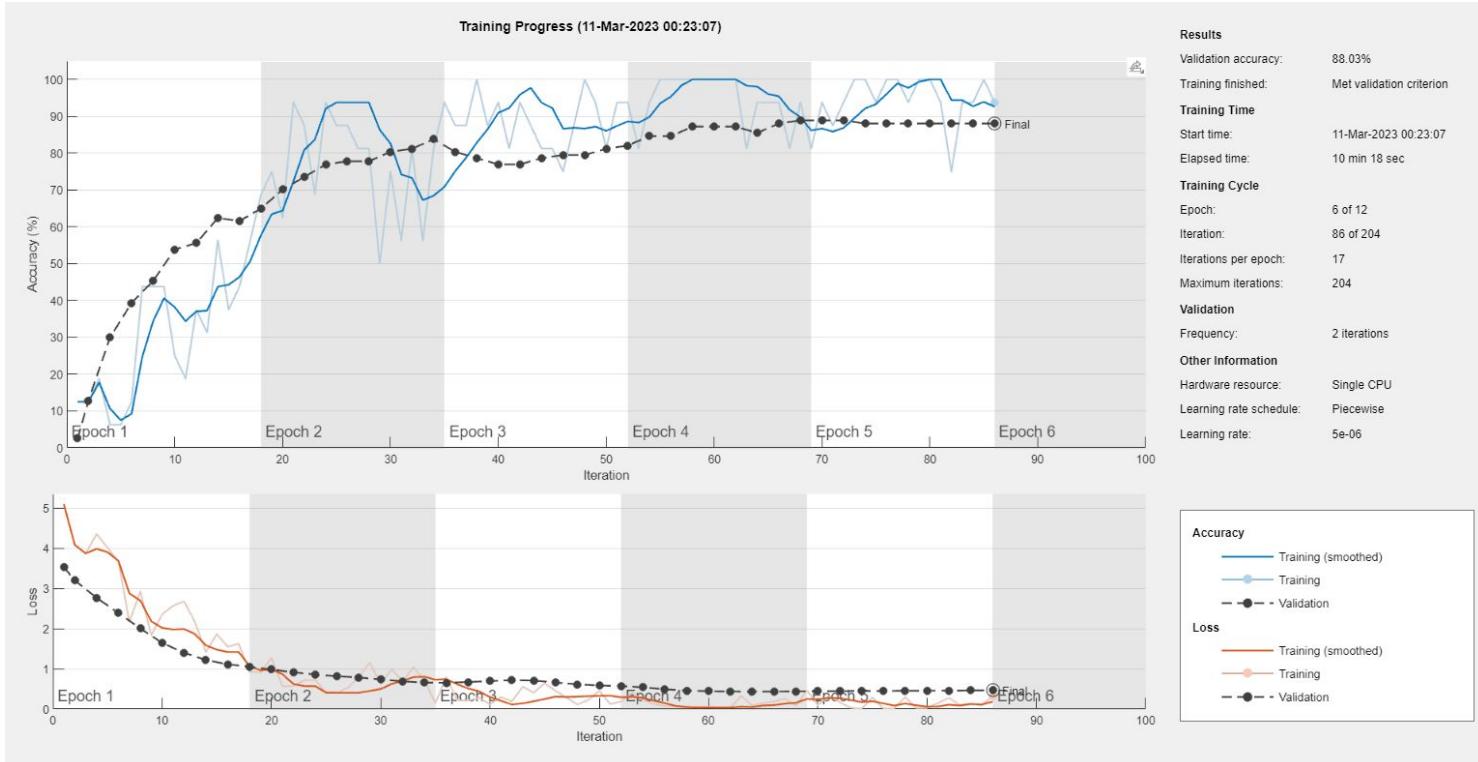
Final Results

- Modified Group 1 Image classes
- Group 1: 284 images
- Group 2: 399 images
- Data used for all 4 CNN architectures
- Only Final fully connected layer modified
- Data Augmentation:
 - Random Rotation (-20,20) degrees,
 - Random X-Y Translation (-3,3) pixels

<u>Old Classes</u>	<u>New Classes</u>
Benson	Benson
Bronco Statue	Bronco Statue
Casa	Dowd
Daly	Graham
Dowd	Library
Graham	Lucas Hall
Kenna	Mission Church
Library	O'Connor
Mission Church	SCDI
O'Connor	Vari Hall

AlexNet: Final Results

Gr. 1 Final Results: Val. Accuracy/Loss



Final Results: Confusion Matrix

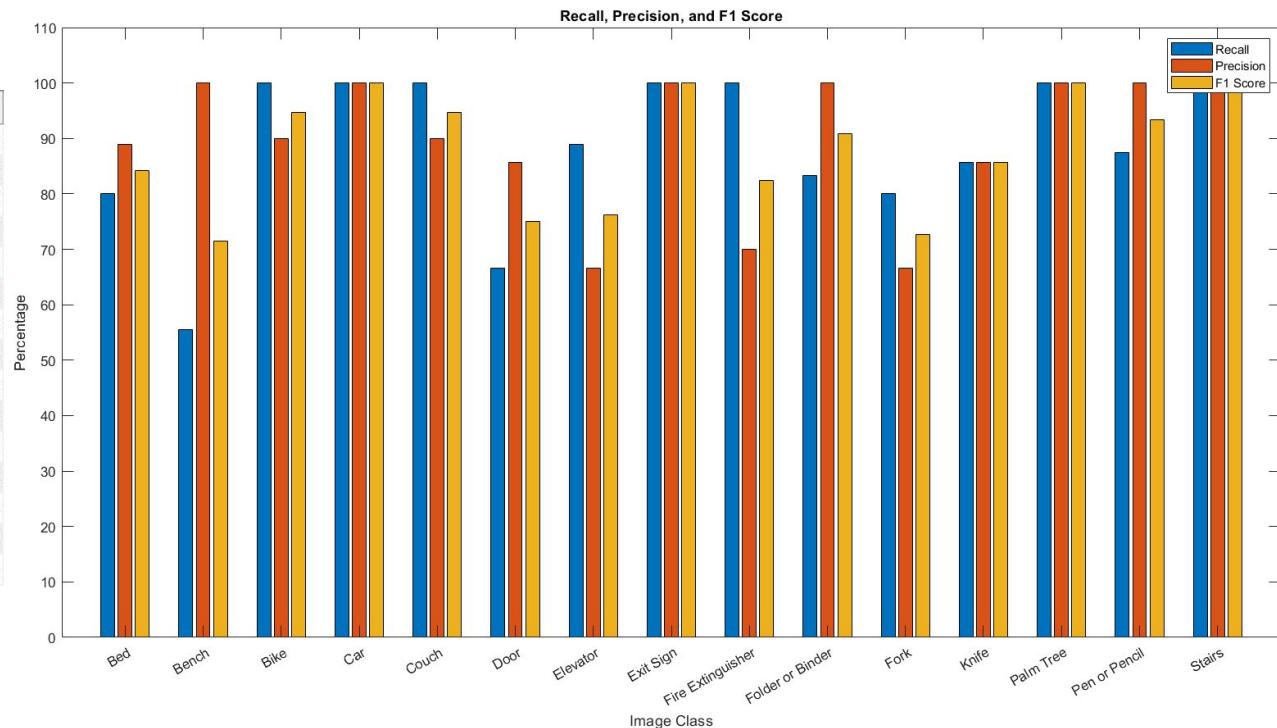
- | | |
|--------------|----------------|
| 1. Bed | 10. Folder/ |
| 2. Bench | Binder |
| 3. Bike | 11. Fork |
| 4. Car | 12. Knife |
| 5. Couch | 13. Palm Tree |
| 6. Door | 14. Pen/Pencil |
| 7. Elevator | 15. Stairs |
| 8. Exit Sign | |
| 9. Fire | |
| Extinguisher | |

Confusion Matrix; Accuracy = 0.88034; Loss = 0.46466															
True Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	8				1	1									
	1	5	1				1								
			9												
				7											
					9										
						6	2								
							8				1				
								6							
									7						
										5	1				
											4	1			
											1	6			
												7			
													7		
															9

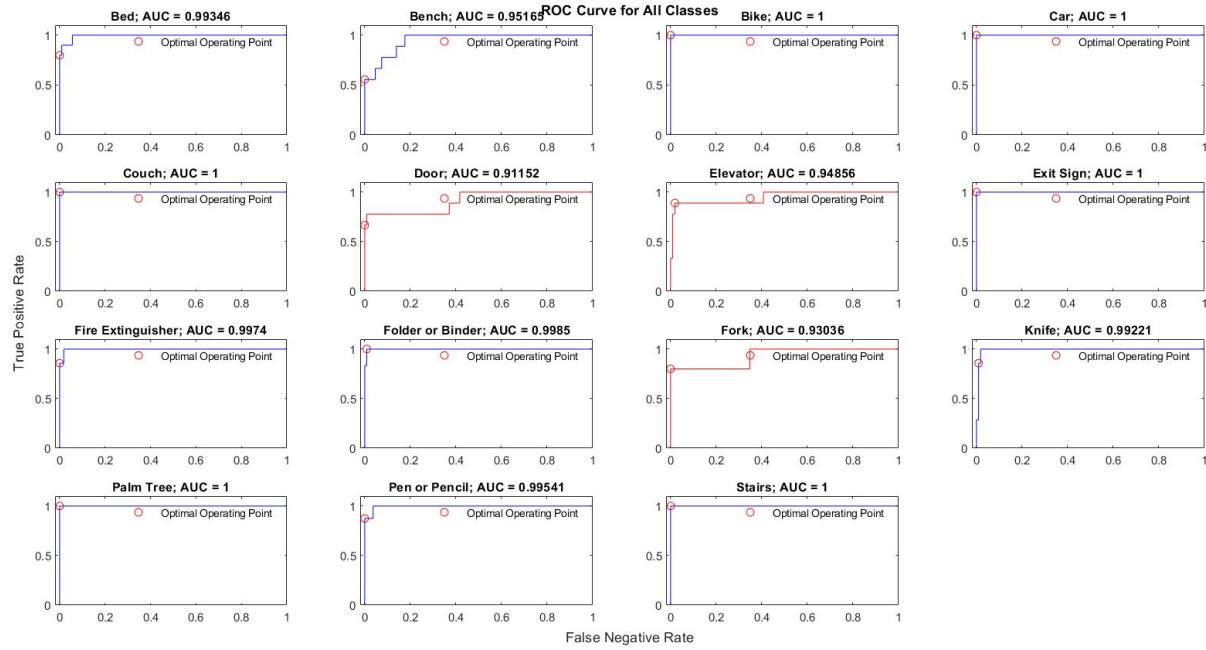
Predicted Class

Final Results: Recall, Precision, F1 Score

	Recall	Precision	F1
Bed	80	88.8889	84.2105
Bench	55.5556	100	71.4286
Bike	100	90	94.7368
Car	100	100	100
Couch	100	90	94.7368
Door	66.6667	85.7143	75
Elevator	88.8889	66.6667	76.1905
Exit Sign	100	100	100
Fire Extinguisher	100	70	82.3529
Folder or Binder	83.3333	100	90.9091
Fork	80	66.6667	72.7273
Knife	85.7143	85.7143	85.7143
Palm Tree	100	100	100
Pen or Pencil	87.5000	100	93.3333
Stairs	100	100	100



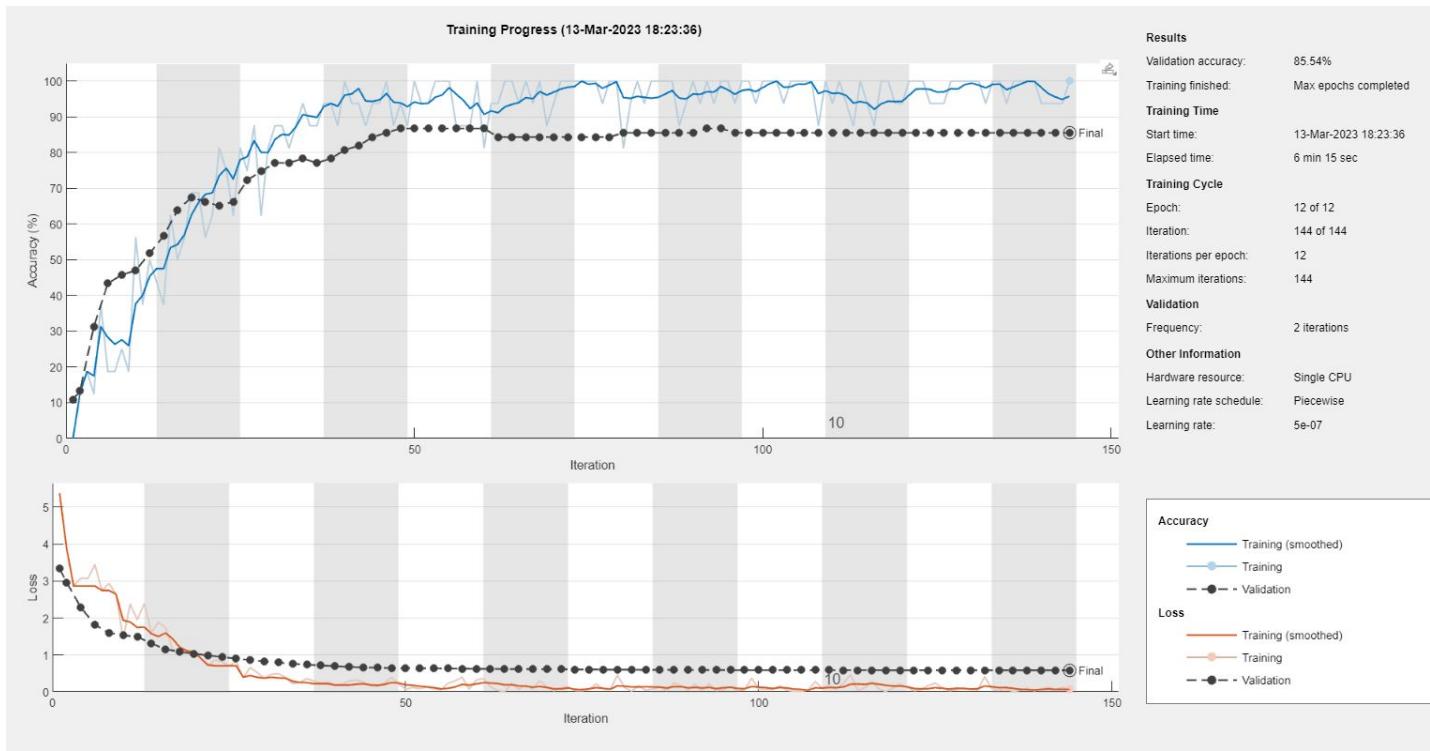
Final Results: ROC Curves



Final Results: Predictions



Gr. 2 Final Results: Val. Accuracy/Loss



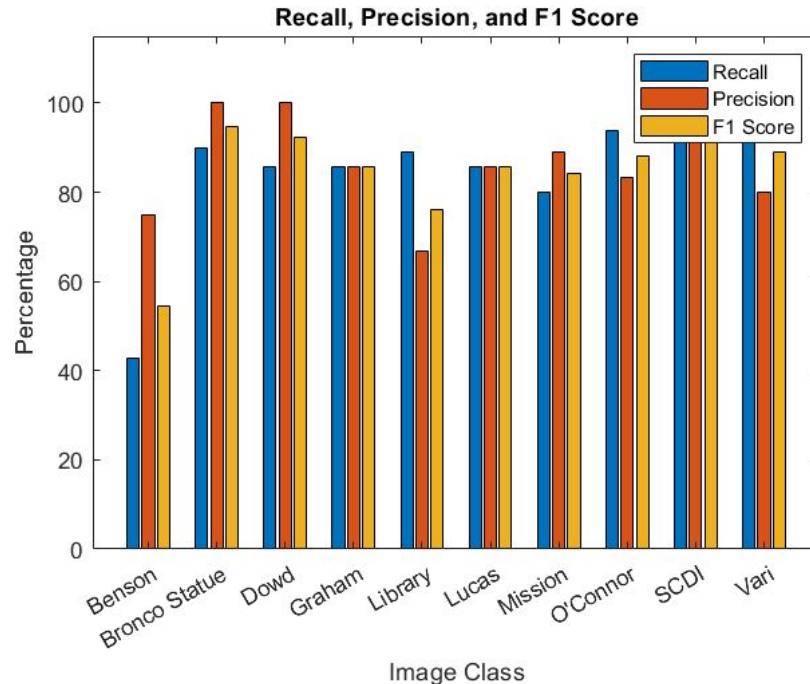
Final Results: Confusion Matrix

1. Benson
 2. Bronco Statue
 3. Casa RLC
 4. Daly Science
 5. Dowd
 6. Graham RLC
 7. Kenna
 8. Library
 9. Mission Church
 10. O'Connor

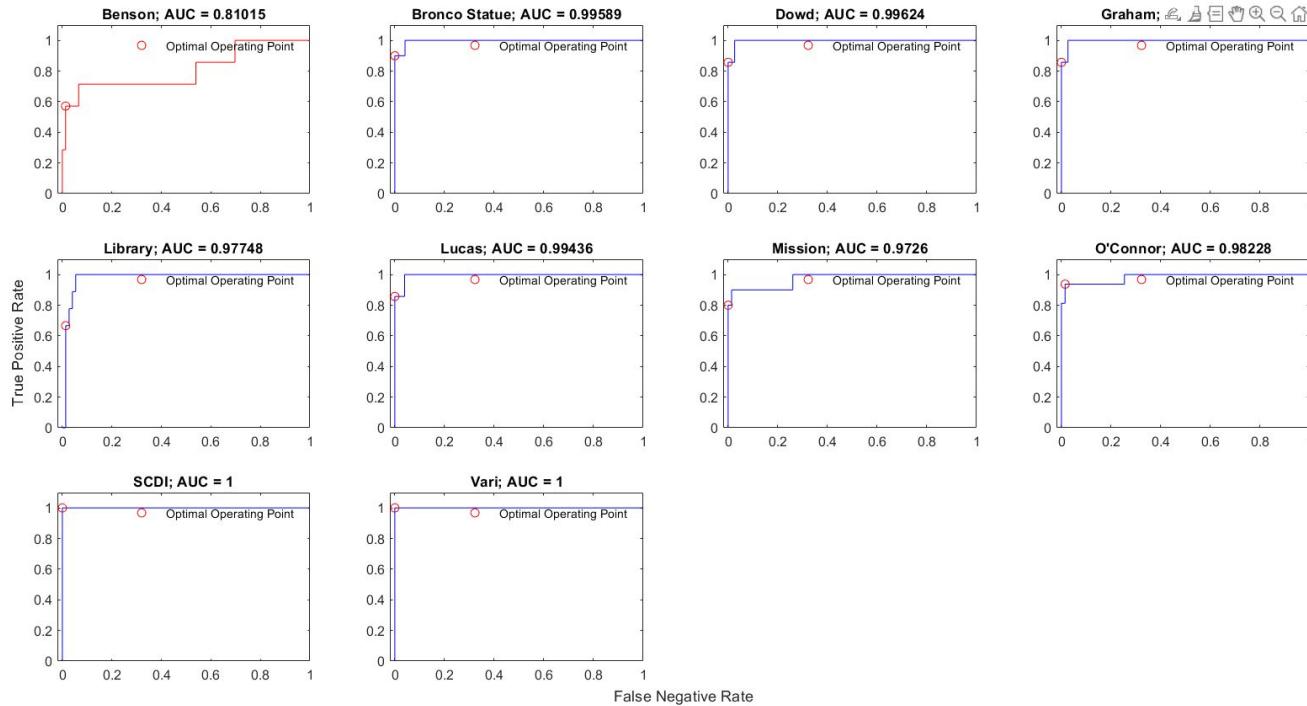
Confusion Matrix; Accuracy = 0.85542; Loss = 0.58543

Final Results: Recall, Precision, F1 Score

	Recall	Precision	F1
Benson	42.8571	75	54.5455
Bronco Statue	90	100	94.7368
Dowd	85.7143	100	92.3077
Graham	85.7143	85.7143	85.7143
Library	88.8889	66.6667	76.1905
Lucas	85.7143	85.7143	85.7143
Mission	80	88.8889	84.2105
O'Connor	93.7500	83.3333	88.2353
SCDI	100	100	100
Vari	100	80	88.8889



Final Results: ROC Curves



Final Results: Predictions

Vari, 99%



O'Connor, 100%



Graham, 99.5%



Library, 60.5%



Lucas, 94.7%



Mission, 57.7%



O'Connor, 100%



Benson, 38.8%



Lucas, 42.3%



O'Connor, 99.9%



O'Connor, 39%



Lucas, 91%



O'Connor, 100%



Benson, 99.9%



Lucas, 60.3%

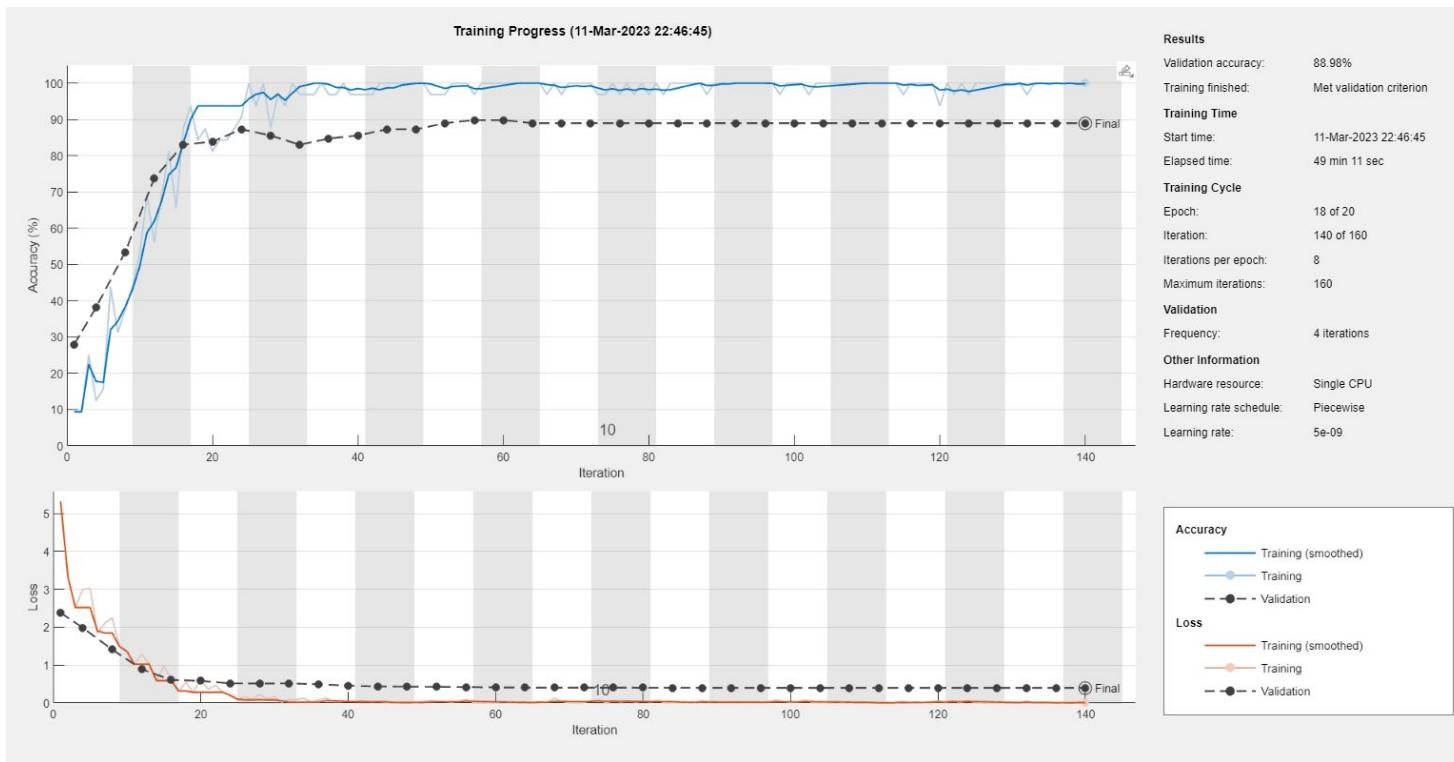


O'Connor, 83.4%



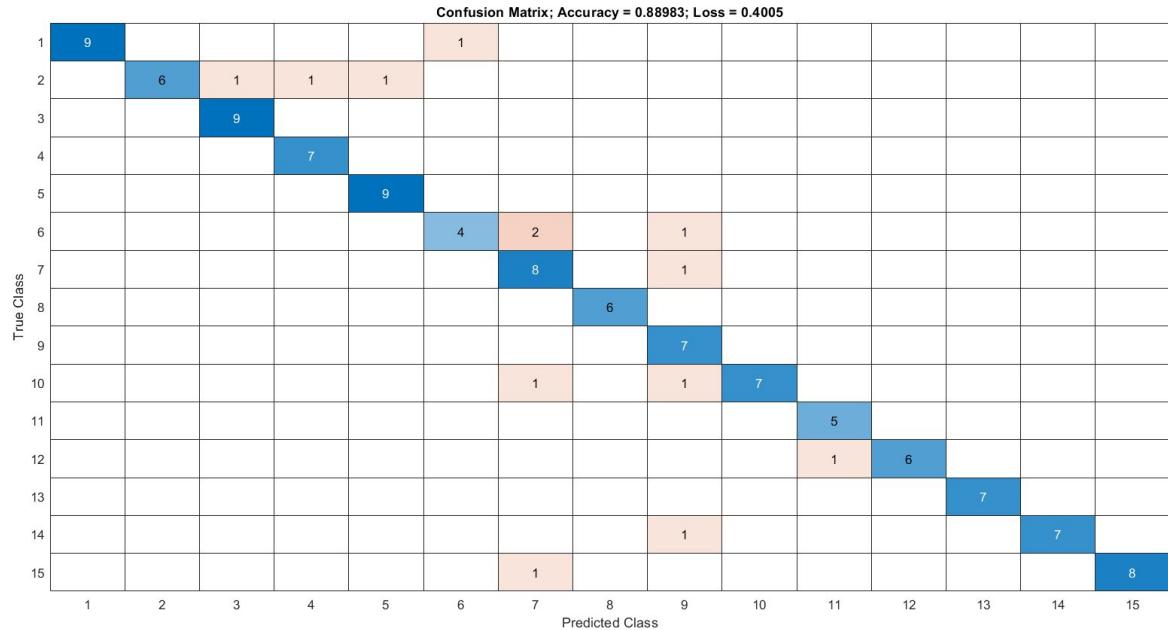
VGG-16: Final Results

Gr. 1 Final Results: Val. Accuracy/Loss



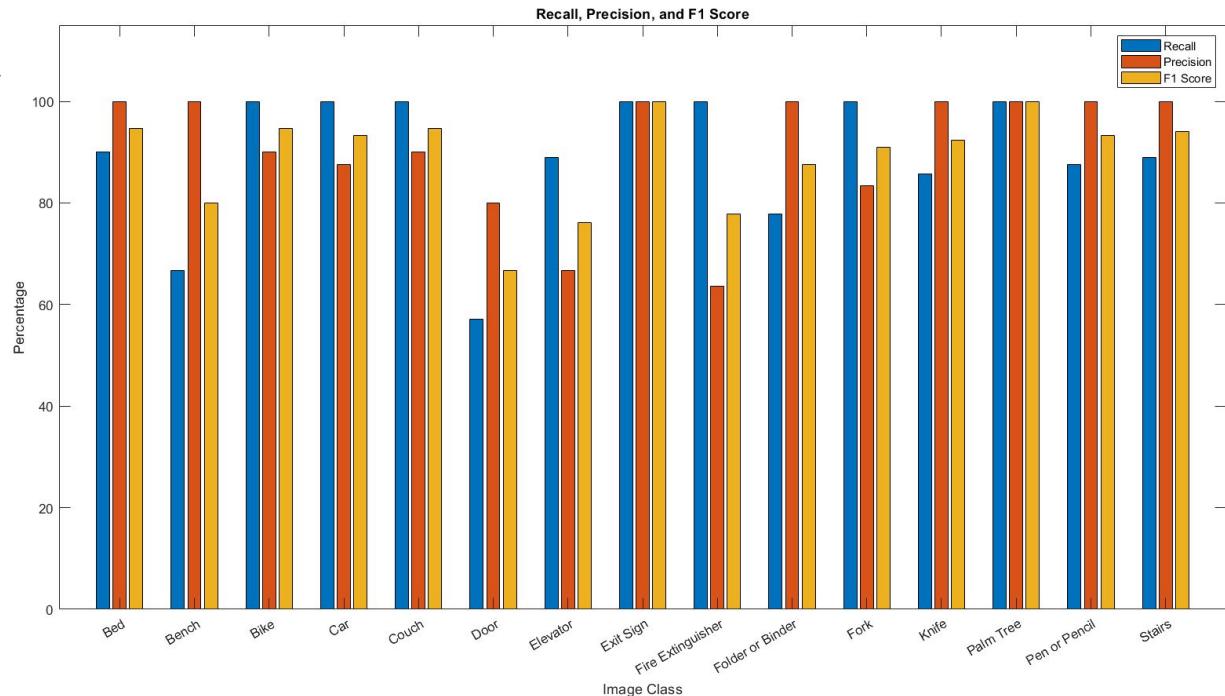
Final Results: Confusion Matrix

- | | |
|--------------|----------------|
| 1. Bed | 10. Folder/ |
| 2. Bench | Binder |
| 3. Bike | 11. Fork |
| 4. Car | 12. Knife |
| 5. Couch | 13. Palm Tree |
| 6. Door | 14. Pen/Pencil |
| 7. Elevator | 15. Stairs |
| 8. Exit Sign | |
| 9. Fire | |
| Extinguisher | |

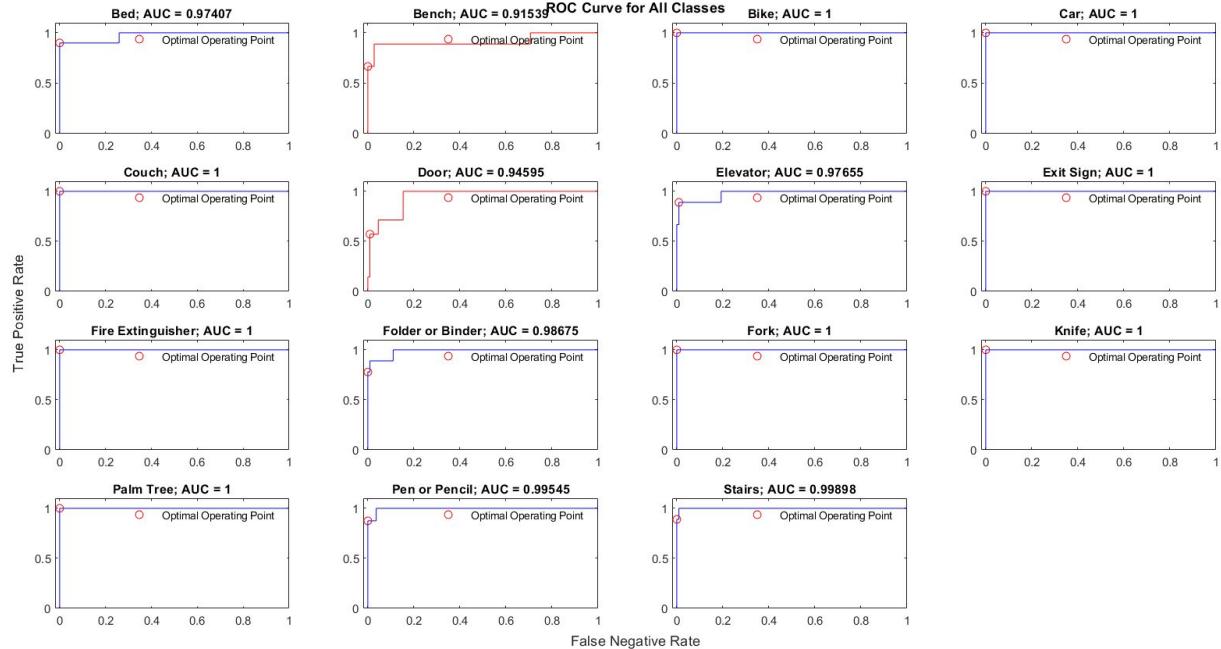


Final Results: Recall, Precision, F1 Score

	Recall	Precision	F1
Bed	90	100	94.7368
Bench	66.6667	100	80
Bike	100	90	94.7368
Car	100	87.5000	93.3333
Couch	100	90	94.7368
Door	57.1429	80	66.6667
Elevator	88.8889	66.6667	76.1905
Exit Sign	100	100	100
Fire Extinguisher	100	63.6364	77.7778
Folder or Binder	77.7778	100	87.5000
Fork	100	83.3333	90.9091
Knife	85.7143	100	92.3077
Palm Tree	100	100	100
Pen or Pencil	87.5000	100	93.3333
Stairs	88.8889	100	94.1176



Final Results: ROC Curves



Final Results: Predictions

Couch, 99.9%



Bike, 100%



Knife, 99.7%



Exit S



Pen or Pencil, 98.4%



Fire Extinguisher, 100%



Folder or Binder, 99.6%



Fork, 50.9%



Elevator, 100%



Fire Extinguisher, 72.5%



Car, 97.5%



Stairs, 96.7%



Exit Sign, 70.7%



Couch, 100%



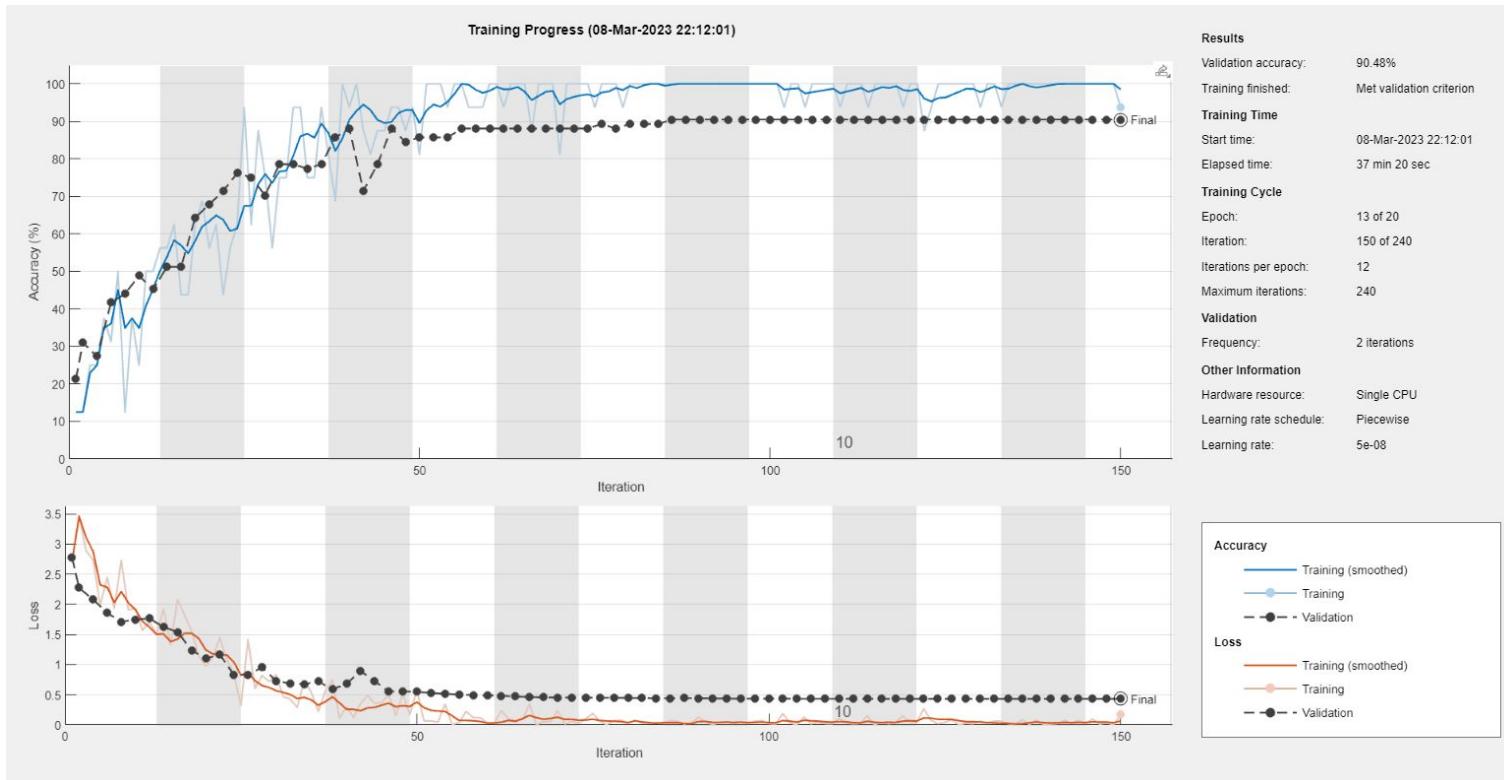
Car, 42.6%



Elevator, 67.8%



Gr. 2 Final Results: Val. Accuracy/Loss



Final Results: Confusion Matrix

1. Benson
2. Bronco Statue
3. Casa RLC
4. Daly Science
5. Dowd
6. Graham RLC
7. Kenna
8. Library
9. Mission Church
10. O'Connor

Confusion Matrix; Accuracy = 0.90476; Loss = 0.43302

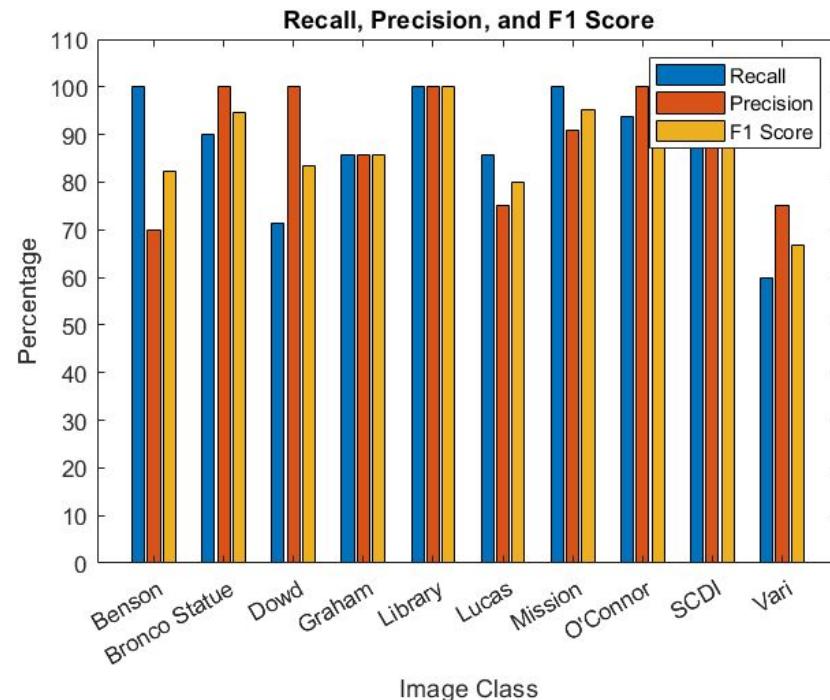
	1	2	3	4	5	6	7	8	9	10
1	7									
2	1	9								
3	1		5	1						
4				6		1				
5					9					
6						6				1
7							10			
8							1	15		
9									6	
10	1					1				3

True Class

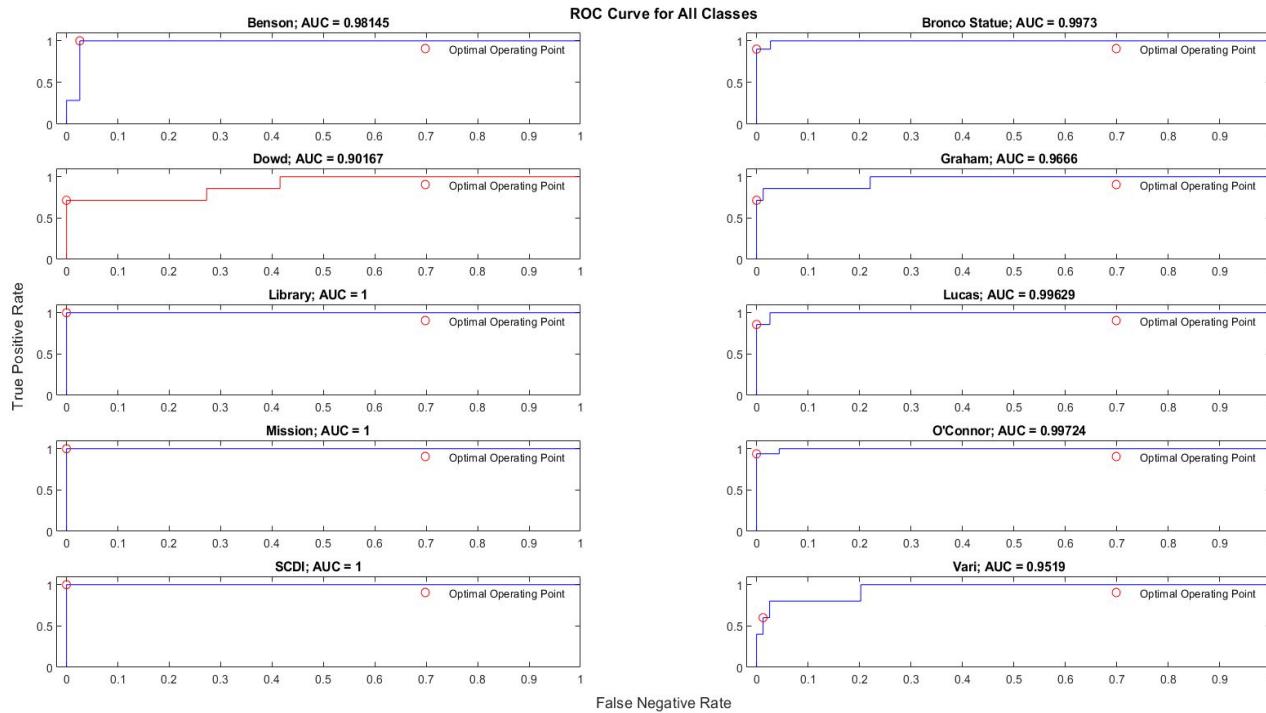
Predicted Class

Final Results: Recall, Precision, F1 Score

	Recall	Precision	F1
Benson	100	70	82.3529
Bronco Statue	90	100	94.7368
Dowd	71.4286	100	83.3333
Graham	85.7143	85.7143	85.7143
Library	100	100	100
Lucas	85.7143	75	80
Mission	100	90.9091	95.2381
O'Connor	93.7500	100	96.7742
SCDI	100	100	100
Vari	60	75	66.6667



Final Results: ROC Curves



Final Results: Predictions

O'Connor, 100%



Bronco Statue, 100%



Mission, 99.3%



Benson, 88.5%



SCDI, 100%



Benson, 98%



Mission, 100%



Graham, 100%



Mission, 100%



Lucas, 91.3%



O'Connor, 99.8%



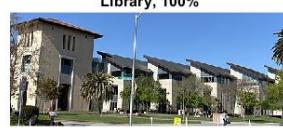
O'Connor, 99.9%



Graham, 100%



Library, 100%



Dowd, 99.9%

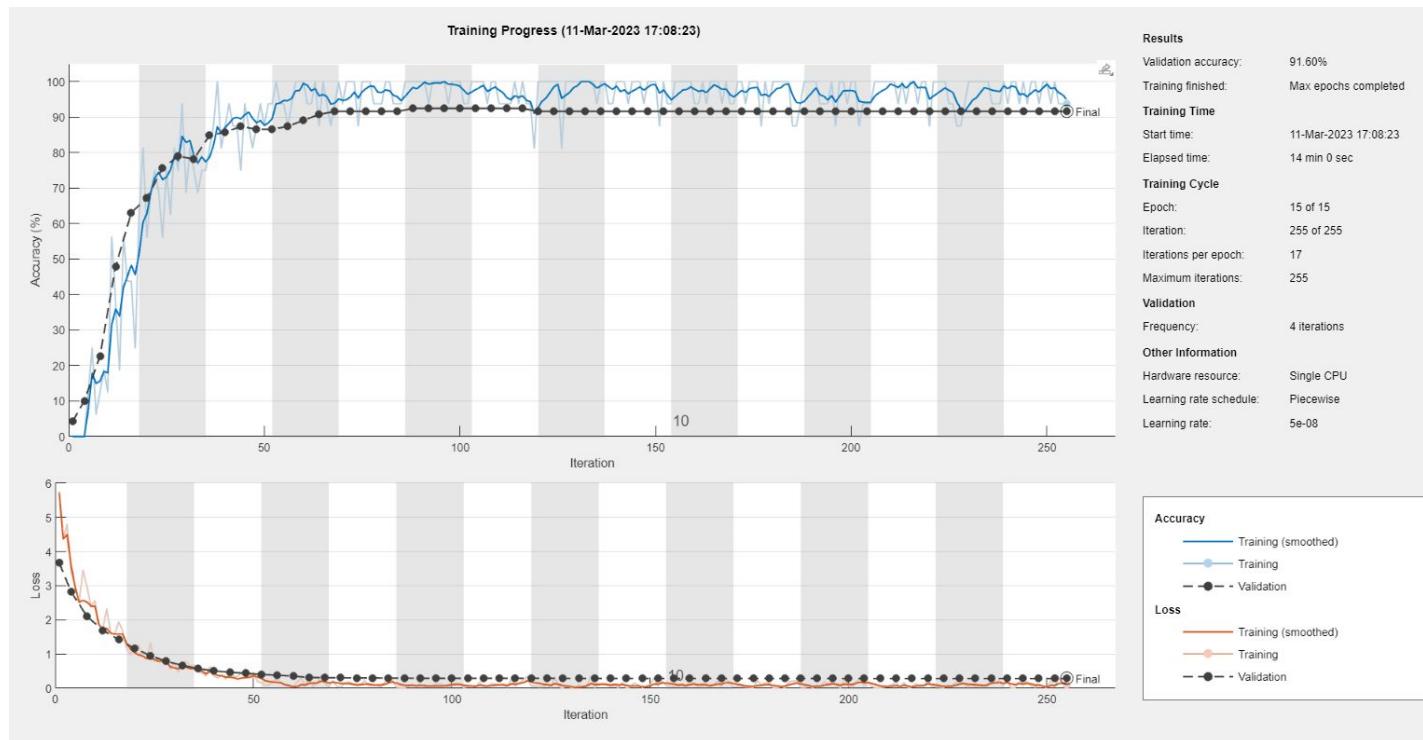


Benson, 91%



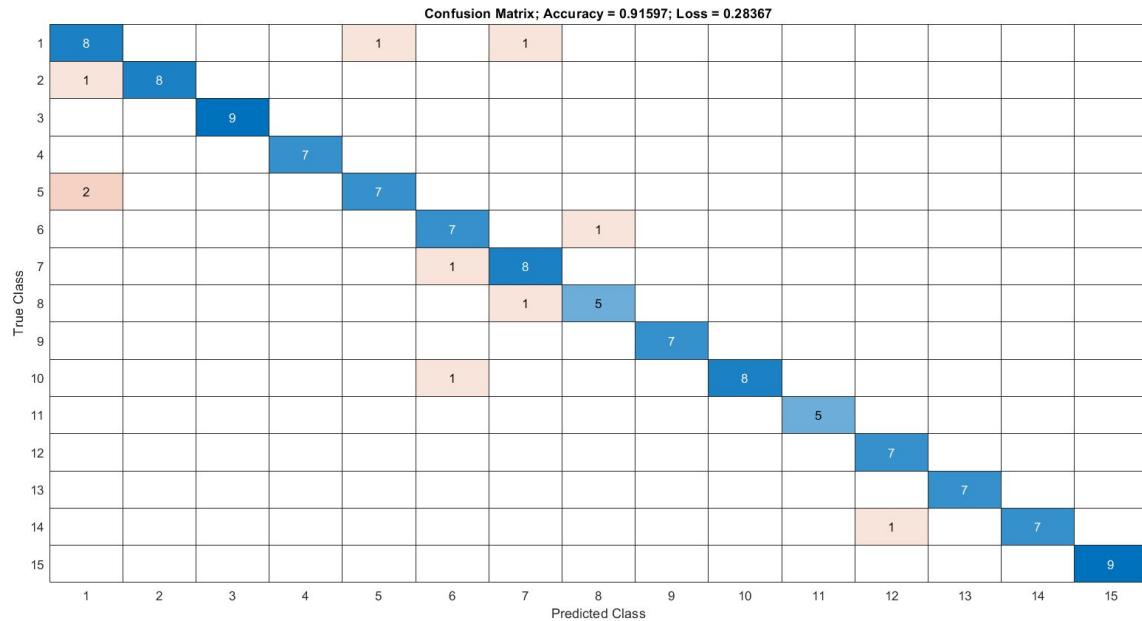
GoogleNet: Final Results

Gr. 1 Final Results: Val. Accuracy/Loss



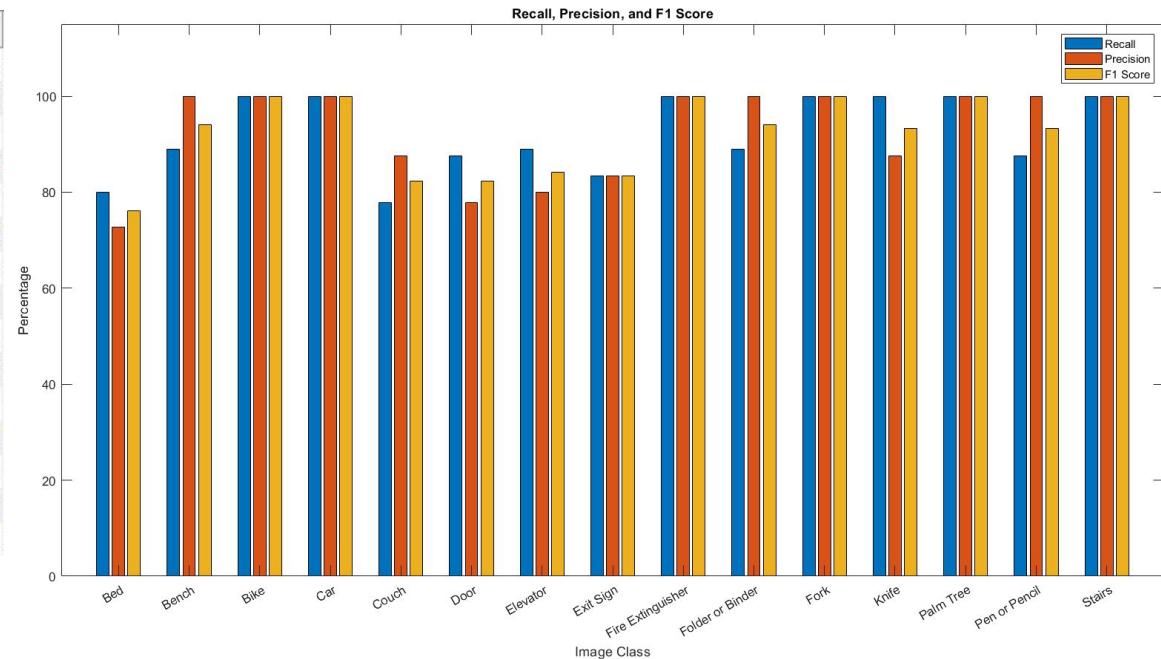
Final Results: Confusion Matrix

- | | | | |
|----|--------------|-----|-------------------|
| 1. | Bed | 10. | Folder/
Binder |
| 2. | Bench | 11. | Fork |
| 3. | Bike | 12. | Knife |
| 4. | Car | 13. | Palm Tree |
| 5. | Couch | 14. | Pen/Pencil |
| 6. | Door | 15. | Stairs |
| 7. | Elevator | | |
| 8. | Exit Sign | | |
| 9. | Fire | | |
| | Extinguisher | | |

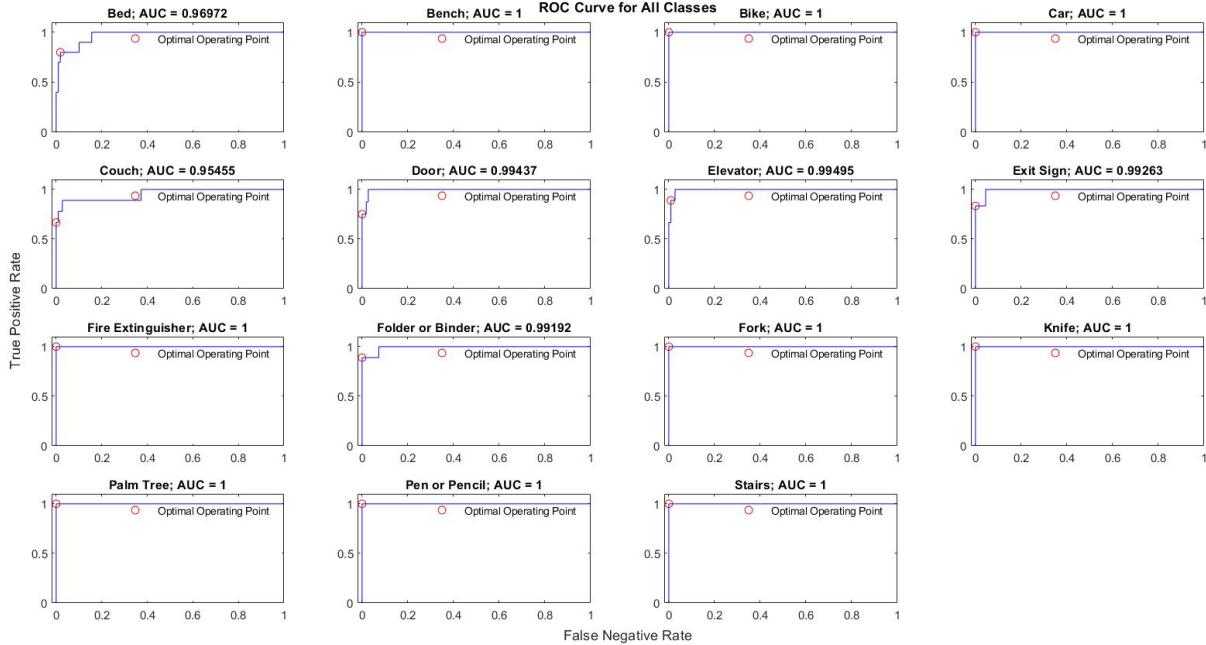


Final Results: Recall, Precision, F1 Score

	Recall	Precision	F1
Bed	80	72.7273	76.1905
Bench	88.8889	100	94.1176
Bike	100	100	100
Car	100	100	100
Couch	77.7778	87.5000	82.3529
Door	87.5000	77.7778	82.3529
Elevator	88.8889	80	84.2105
Exit Sign	83.3333	83.3333	83.3333
Fire Extinguisher	100	100	100
Folder or Binder	88.8889	100	94.1176
Fork	100	100	100
Knife	100	87.5000	93.3333
Palm Tree	100	100	100
Pen or Pencil	87.5000	100	93.3333
Stairs	100	100	100



Final Results: ROC Curves



Final Results: Predictions

Fork, 98.4%



Stairs, 98.3%



Fork, 98.2%



Bench, 67.4%



Exit Sign, 100%



Knife, 99.6%



Exit Sign, 100%



Stairs, 98.9%



Stairs, 93.4%



Pen or Pencil, 98.5%



Bed, 87.5%



Fire Extinguisher, 98.9%



Folder or Binder, 88.4%



Elevator, 44.9%



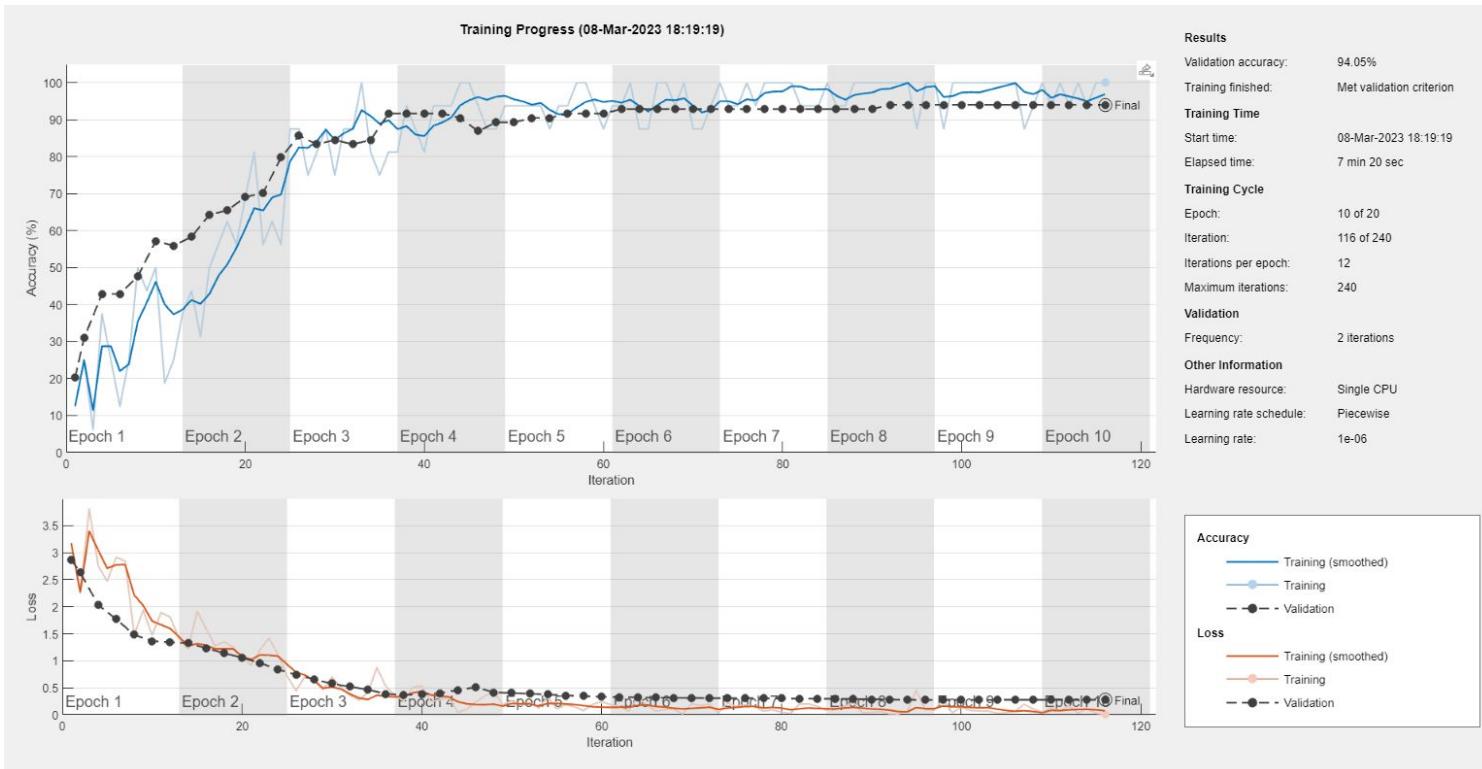
Stairs, 82.2%



Couch, 75.4%



Gr. 2 Final Results: Val. Accuracy/Loss



Final Results: Confusion Matrix

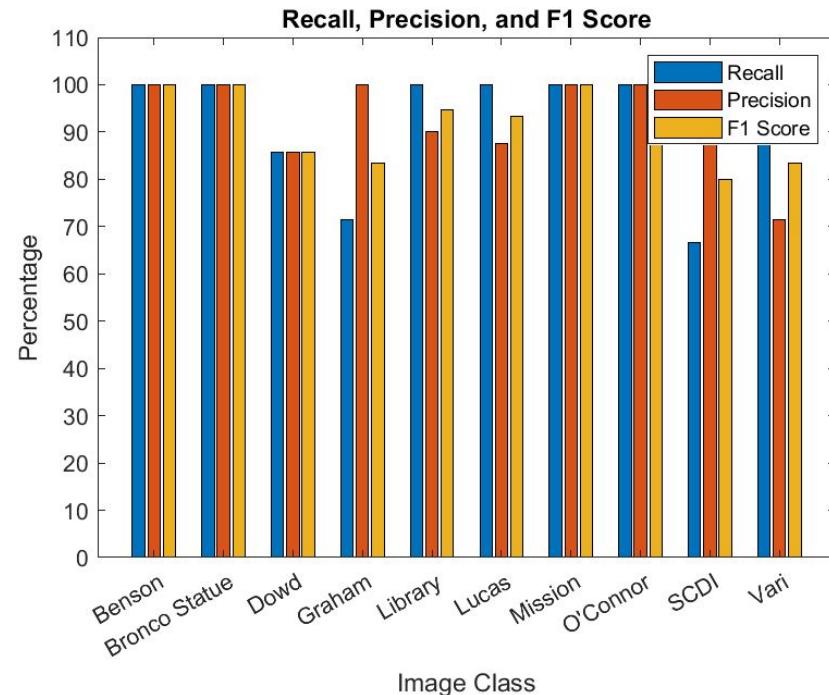
1. Benson
 2. Bronco Statue
 3. Casa RLC
 4. Daly Science
 5. Dowd
 6. Graham RLC
 7. Kenna
 8. Library
 9. Mission Church
 10. O'Connor

Confusion Matrix; Accuracy = 0.94048; Loss = 0.43488

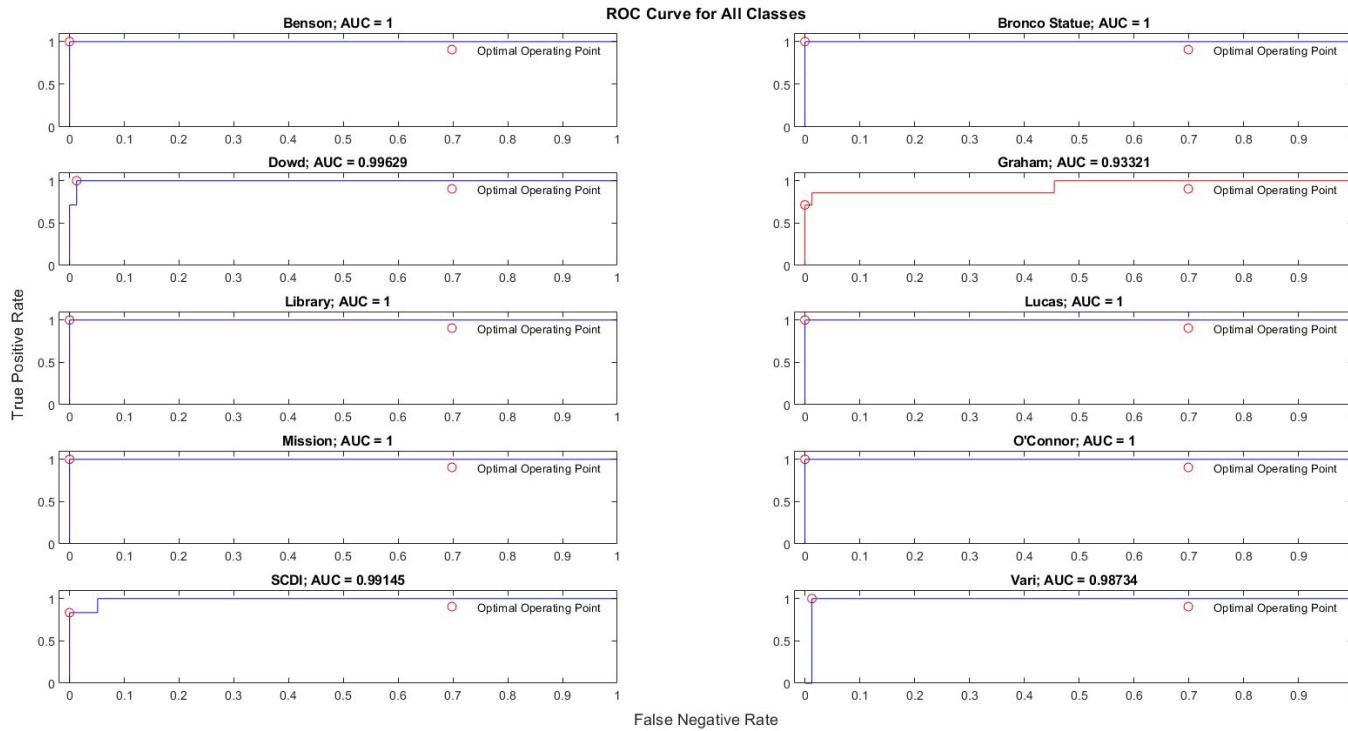
True Class	1	2	3	4	5	6	7	8	9	10
1	7									
2		10								
3			6				1			
4			1	5						1
5					9					
6						7				
7							10			
8								16		
9					1				4	1
10										5
Predicted Class	1	2	3	4	5	6	7	8	9	10

Final Results: Recall, Precision, F1 Score

	Recall	Precision	F1
Benson	100	100	100
Bronco Statue	100	100	100
Dowd	85.7143	85.7143	85.7143
Graham	71.4286	100	83.3333
Library	100	90	94.7368
Lucas	100	87.5000	93.3333
Mission	100	100	100
O'Connor	100	100	100
SCDI	66.6667	100	80
Vari	100	71.4286	83.3333



Final Results: ROC Curves



Final Results: Predictions

Dowd, 49.5%



Vari, 94.6%



Library, 83.3%



Dowd, 49%



O'Connor, 100%



Vari, 82.7%



Benson, 99.9%



Bronco Statue, 100%



Mission, 100%



Dowd, 81.5%



Bronco Statue, 100%



Library, 99%



Mission, 100%



Bronco Statue, 100%



Mission, 99.8%

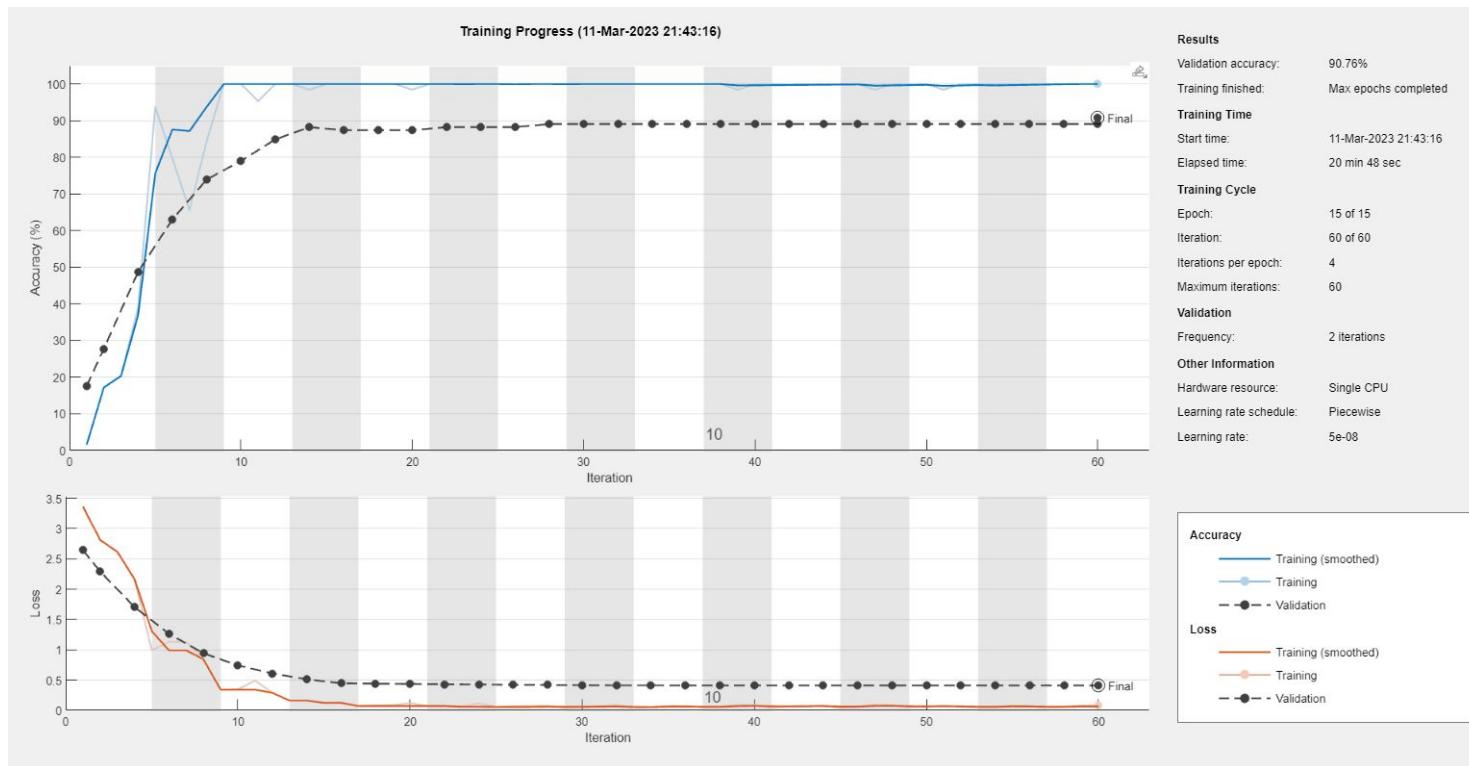


Bronco Statue, 100%



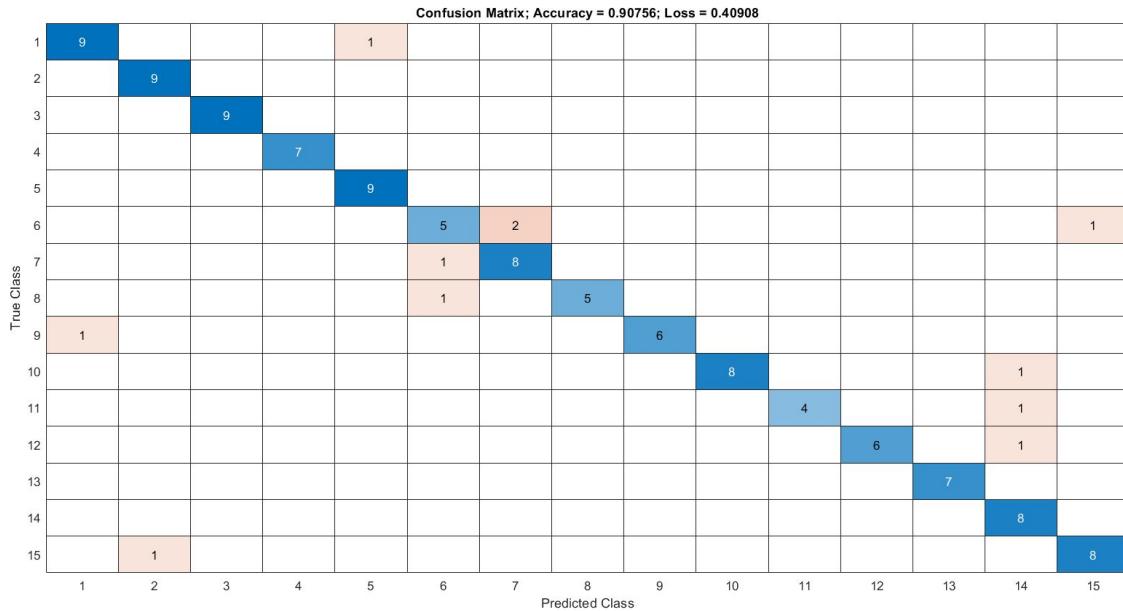
ResNet50: Final Results

Gr. 1 Final Results: Val. Accuracy/Loss



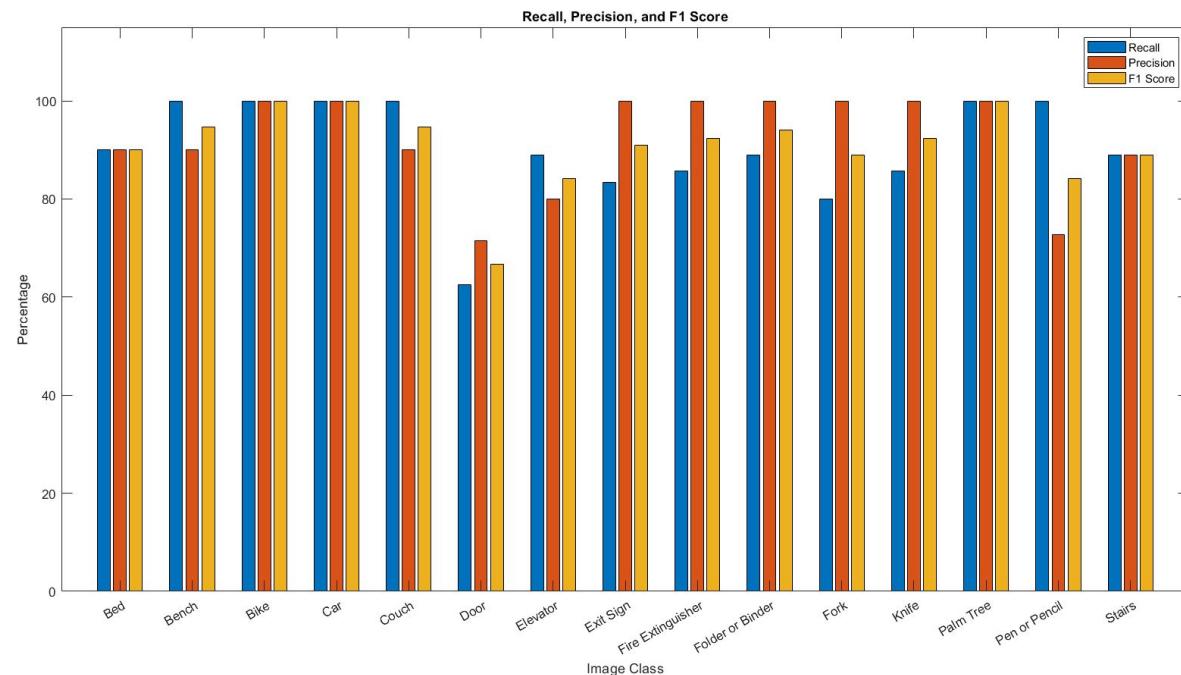
Final Results: Confusion Matrix

- | | | | |
|----|--------------|-----|------------|
| 1. | Bed | 10. | Folder/ |
| 2. | Bench | | Binder |
| 3. | Bike | 11. | Fork |
| 4. | Car | 12. | Knife |
| 5. | Couch | 13. | Palm Tree |
| 6. | Door | 14. | Pen/Pencil |
| 7. | Elevator | 15. | Stairs |
| 8. | Exit Sign | | |
| 9. | Fire | | |
| | Extinguisher | | |

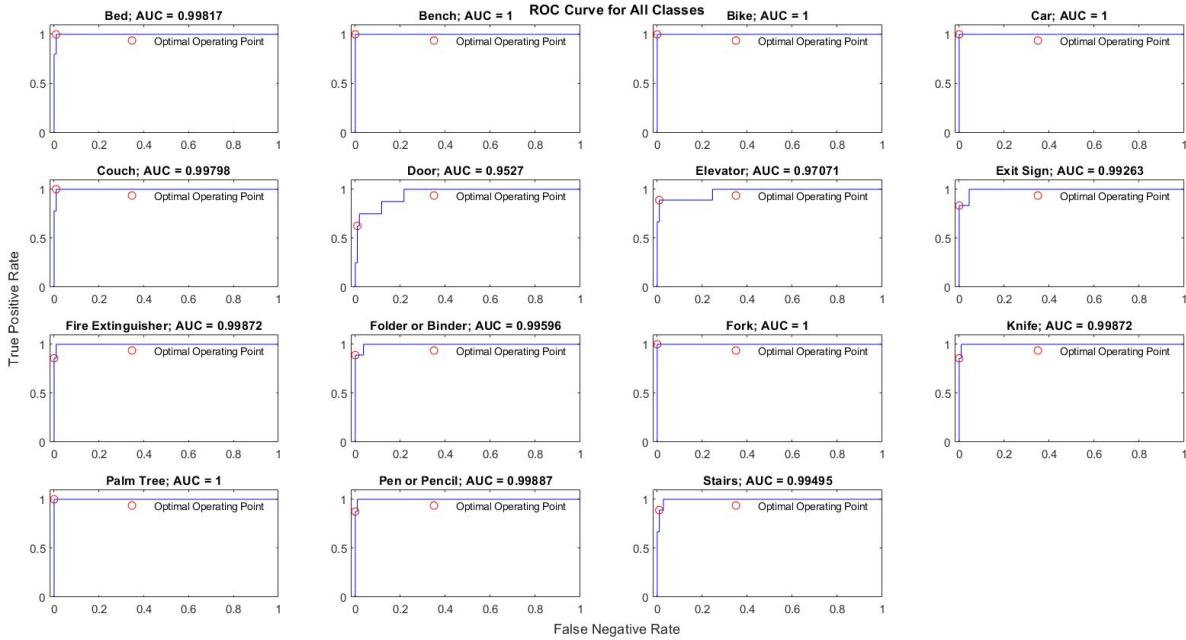


Final Results: Recall, Precision, F1 Score

	Recall	Precision	F1
Bed	90	90	90
Bench	100	90	94.7368
Bike	100	100	100
Car	100	100	100
Couch	100	90	94.7368
Door	62.5000	71.4286	66.6667
Elevator	88.8889	80	84.2105
Exit Sign	83.3333	100	90.9091
Fire Extinguisher	85.7143	100	92.3077
Folder or Binder	88.8889	100	94.1176
Fork	80	100	88.8889
Knife	85.7143	100	92.3077
Palm Tree	100	100	100
Pen or Pencil	100	72.7273	84.2105
Stairs	88.8889	88.8889	88.8889



Final Results: ROC Curves



Final Results: Predictions

Door, 33.7%



Elevator, 100%



Bench, 96.6%



Exit Sign, 44.6%



Bench, 96.4%



Pen or Pencil, 97.5%



Bed, 37.5%



Folder or Binder, 54.6%



Folder or Binder, 21.4%



Bench, 56.4%



Stairs, 57.1%



Elevator, 92.8%



Pen or Pencil, 51.6%



Elevator, 48.3%



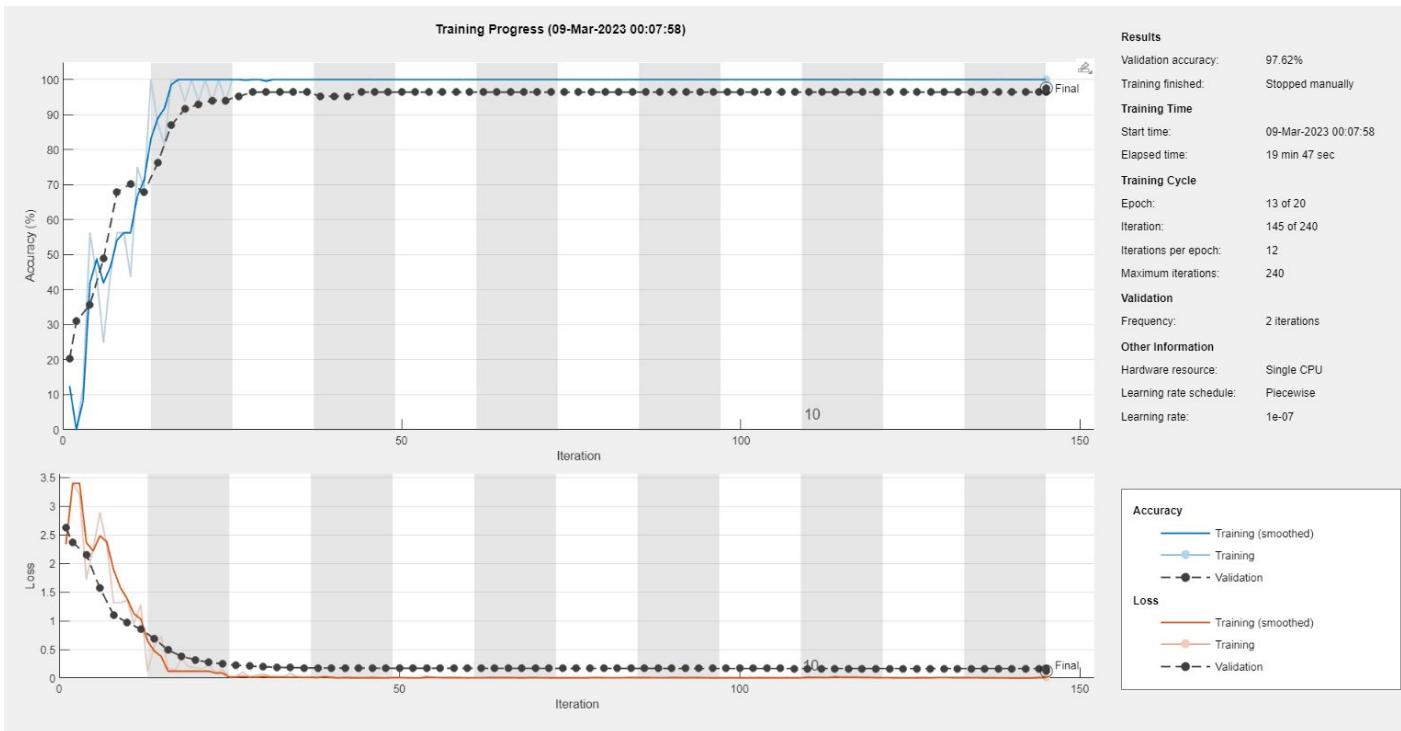
Bed, 48.1%



Stairs, 61.5%

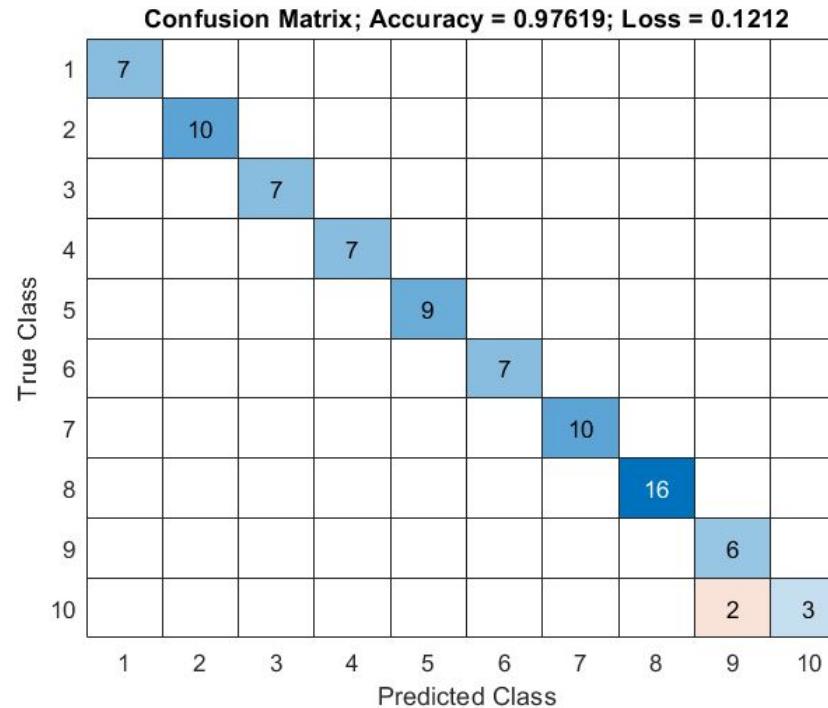


Gr. 2 Final Results: Val. Accuracy/Loss



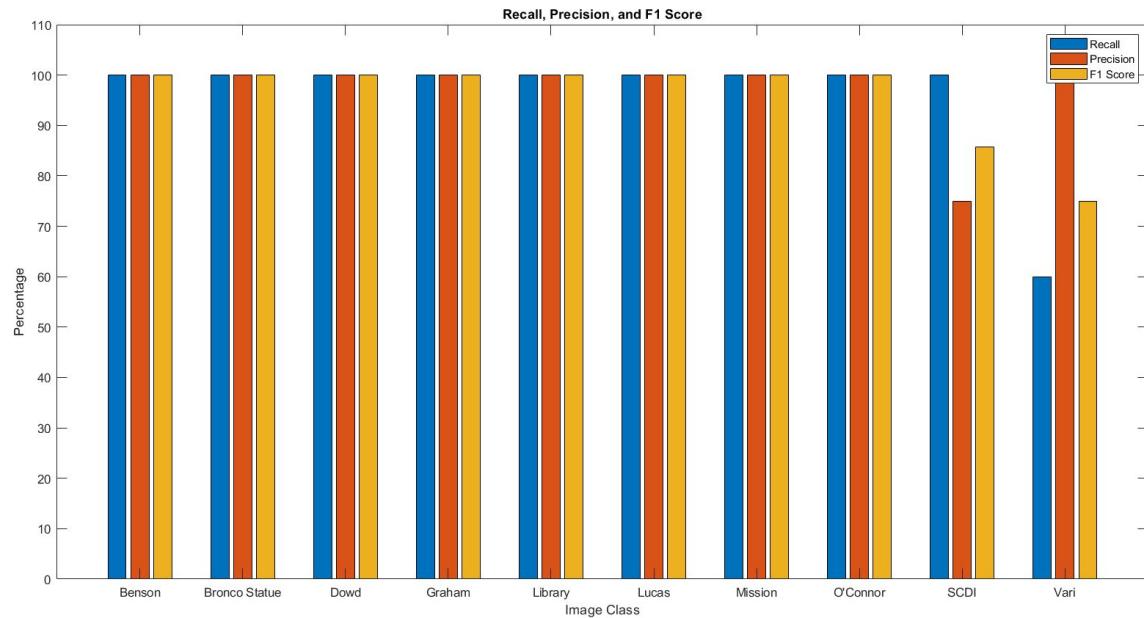
Final Results: Confusion Matrix

1. Benson
2. Bronco Statue
3. Casa RLC
4. Daly Science
5. Dowd
6. Graham RLC
7. Kenna
8. Library
9. Mission Church
10. O'Connor

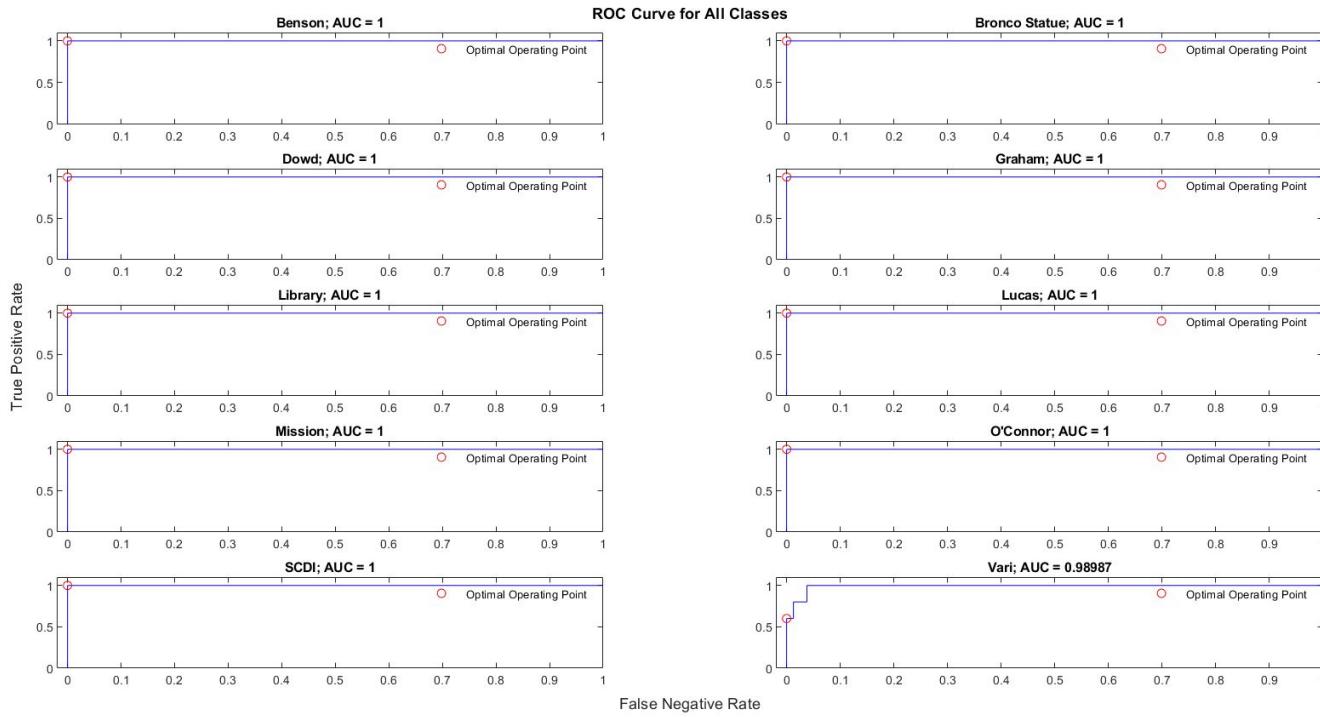


Final Results: Recall, Precision, F1 Score

	Recall	Precision	F1
Benson	100	100	100
Bronco Statue	100	100	100
Dowd	100	100	100
Graham	100	100	100
Library	100	100	100
Lucas	100	100	100
Mission	100	100	100
O'Connor	100	100	100
SCDI	100	75	85.7143
Vari	60	100	75



Final Results: ROC Curves



Final Results: Predictions

Lucas, 100%



O'Connor, 100%



Dowd, 87.5%



Bronco Statue, 100%



Dowd, 89.7%



O'Connor, 100%



SCDI, 73.9%



Benson, 99.5%



Benson, 42.1%



Graham, 99.8%



Dowd, 94.3%



O'Connor, 87.9%



Bronco Statue, 100%



Lucas, 92.3%



Dowd, 99.2%



Benson, 99.9%



Accuracy and Loss Comparison: Group 1

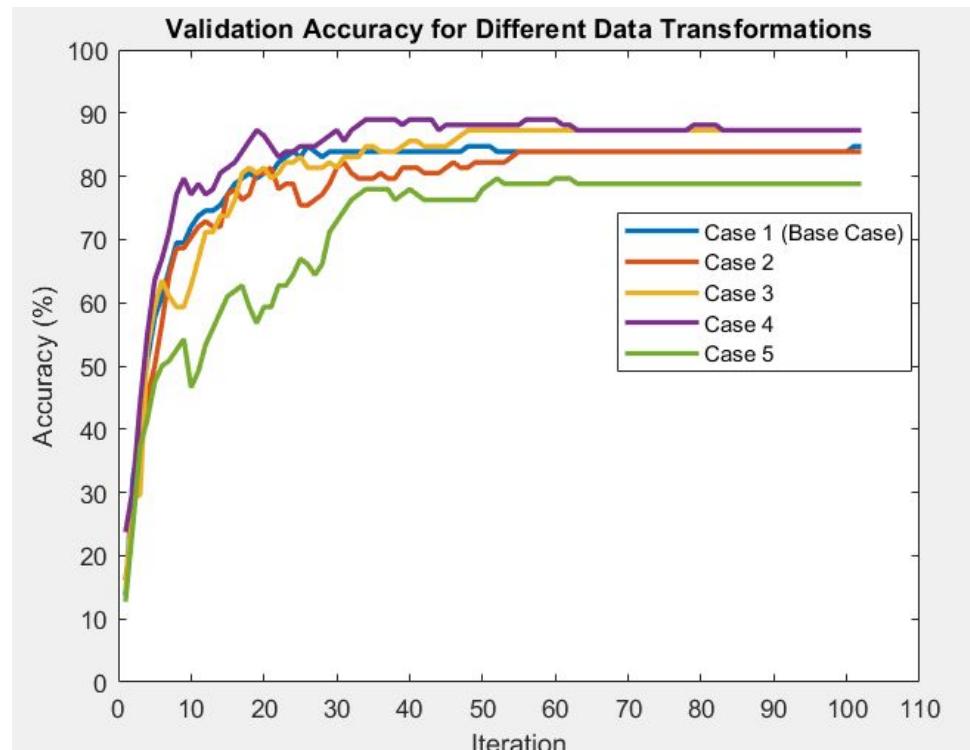
Group 1				
<u>Initial Results</u>				
	AlexNet	VGG-16	GoogleNet	ResNet50
Accuracy	83.33%	-	-	-
Loss	0.6094	-	-	-
<u>Final Results without Data Augmentation</u>				
Accuracy	84.75%	85.60%	86%	89.83%
Loss	0.4692	0.3683	0.3765	0.4219
<u>Final Results with Data Augmentation</u>				
Accuracy	88.03%	88.98%	91.60%	90.76%
Loss	0.4647	0.4005	0.2837	0.4091

Accuracy and Loss Comparison: Group 2

Group 2				
<u>Initial Results</u>				
	<u>AlexNet</u>	<u>VGG-16</u>	<u>GoogleNet</u>	<u>ResNet50</u>
Accuracy	65.39%	-	-	-
Loss	1.0004	-	-	-
<u>Final Results without Data Augmentation</u>				
Accuracy	83.13%	86.75%	80.72%	83.13%
Loss	0.7862	0.66354	0.698	0.6985
<u>Final Results with Data Augmentation</u>				
Accuracy	85.54%	90.48%	94.05%	97.62%
Loss	0.5854	0.433	0.4349	0.1212

Testing Different Data Augmentations

- AlexNet, Group 1, Update Dataset
- Case 1 (Base Case): 84.75%
 - No augmentation
- Case 2: 83.90%
 - Random horizontal and vertical reflections
- Case 3: 87.29%
 - Random horizontal and vertical reflections
 - Random rotations up to 20 degrees
- Case 4: 87.29%
 - Random horizontal and vertical translations up to 50 pixels
- Case 5: 78.81%
 - Random horizontal and vertical shearing
 - Random scaling between 0.5 and 2



Training Time Comparison

	Training Time/Epoch (s)	# parameters (millions)
<u>AlexNet</u>	124	62.3
<u>VGG-16</u>	160	138
<u>GoogleNet</u>	56	6.7977
<u>Resnet50</u>	83	23

Conclusion

- Adding more high-quality data to the dataset improves performance
- Data augmentation is a crucial component to increase accuracy and avoid overfitting
- Other newer CNN architectures are able to give better results than AlexNet, and in less time
- Future work: combine Group 1 and 2 into single dataset



Thank You