UNCC Campus Building Image Recognition Project

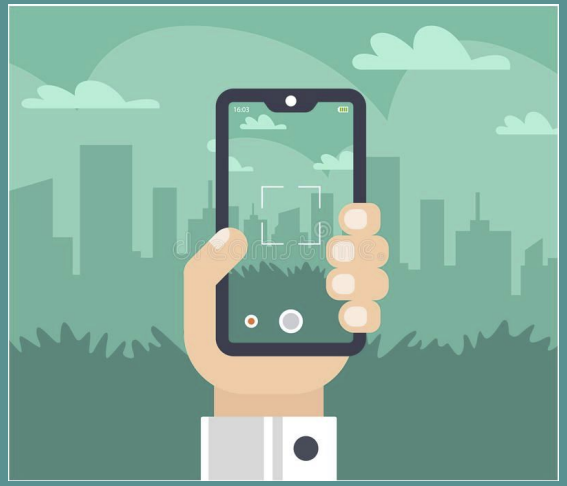
Group 2
Madison Melton
John Taylor
Prasheeth Venkat Kumar
Nathan Williams

Intro and Team Member Contribution

We have collected a sizeable dataset consisting of images of numerous buildings across the UNCC campus. Using this dataset, we have trained a model in order to accurately identify these buildings.

- Madison Melton (Graduate): Original Team Formation, Proposal, Research, Data Collection, Reporting, and Presentation
- John Taylor (Graduate): Project Idea, Proposal, Research, Data Collection, Reporting, and Presentation
- Prasheeth Venkat Kumar (Undergraduate): Proposal, Research, Data Collection, Annotation, Modeling/Training, Reporting, and Presentation
- Nathan Williams (Undergraduate): Proposal, Research, Data Collection, Annotation, Modeling/Training, Reporting, and Presentation

Problem and Motivation



Goal of Project

- Using an image taken by a user on the UNCC campus as input, train a model to accurately predict what building on campus is most likely the subject of the image

Motivation

- To our knowledge, no dataset exists with images of buildings across UNCC
- The scope is confined to very specific area (UNCC) but the model, given it's accuracy, could be used on other campuses or places where numerous buildings need to be identified (large cities?)
- Fun to use a dataset we created / our project relates to our campus

How could this model be used going forward?

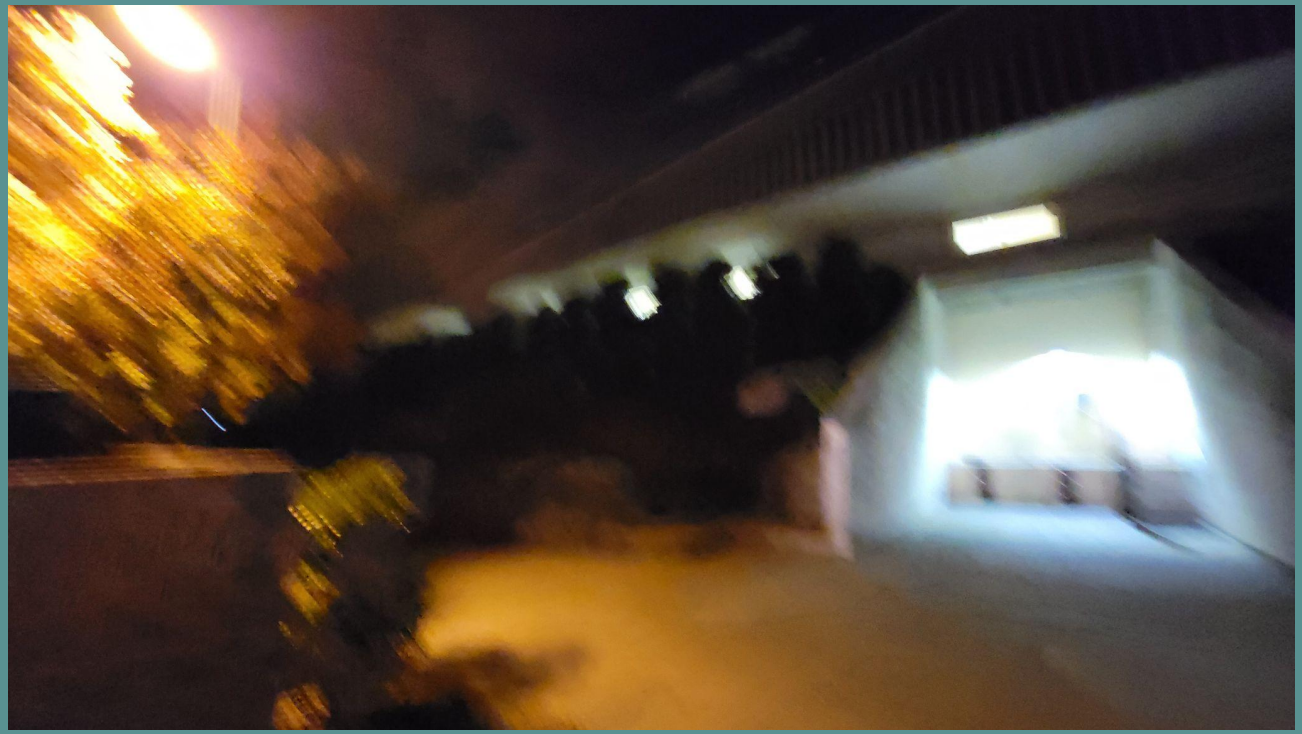
- Self guided tours / Navigation / Etc...

Img Sources: https://pd.charlotte.edu/licensure
https://thumbs.dreamstime.com/b/hand-phone-takes-pictures-city-hand-phone-takes-pictures-city-green-illustration-148568607.jpg

Overview of Dataset

- Over 2700 images, taken during day and night, covering 12 buildings across campus
- Mainly used significant buildings (UREC/Atkins/Etc...)
- Many angles, different lighting scenarios, day/night images
- Annotated
- # of samples in training/test split: 2176 / 544 (80% / 20%)

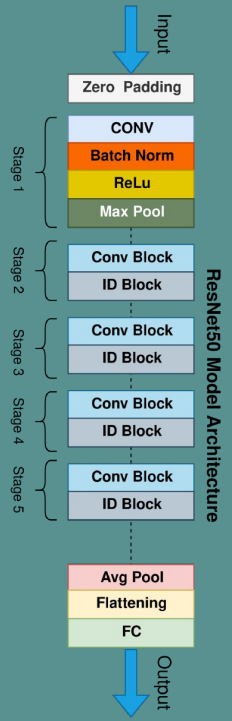
Guess the building!




Belk Gym
Belk Gym
Roy Fielding Aquatic Center
Belk Gymnasium
Lot 7

Methodology: Pre-trained ResNet50 Model

- Research about building recognition, often complex
  - Relatively small dataset
  - Handcrafted features vs Neural Network
  - Building CNN from scratch vs Pre-trained model
  - P. Bezak, "Building recognition system based on deep learning," 2016
- Experimenting:
  - Different Pre-trained models
  - Data augmentation
  - Vision transformers
- Chosen Solution:
  - Pre-trained ResNet50 Model
    - Trained and tuned last layer
- Hyperparameters used:
  - Learning Rate: 0.0001
  - Number of epochs: 15
  - Image Size: (224 x 224)
  - Data Augmentation: Horizontal Flip
  - Data Normalization: Zero Centered
- Loss function: Cross entropy
  - Multiclass image classification
  - Better for non-linear applications
- Optimizer: Adam vs SGD
  - Computationally efficient

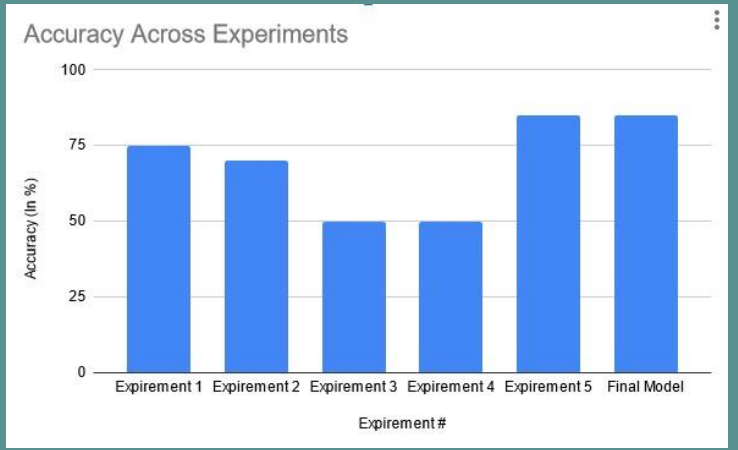


Results

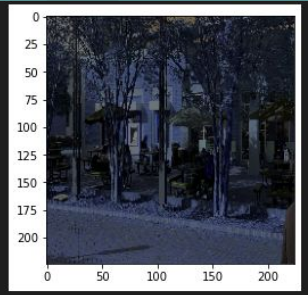
Unless otherwise stated, the following hyperparameters were used in training:

- Learning Rate: 0.0001
- Number of epochs: 15
- Image Size: (224 x 224)
- Data Augmentation: Horizontal Flip
- Data Normalization: Zero Centered

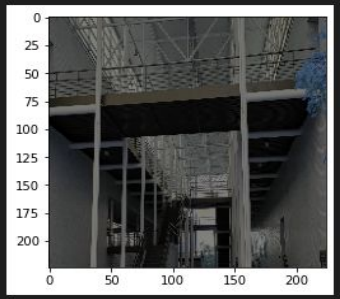
- Experiment 1: Pre-trained GoogleNet
  - Accuracy on test set: ~75%
- Experiment 2: Pre-trained VGG16 model
  - Accuracy on test set: ~67.5%
- Experiment 3: Pre-trained ResNet50 with vision transforming + self attention:
  - Accuracy on test set: ~50%
  - (96 x 96) image size used to reduce CUDA memory needed
  - Learning rate of 0.00005 used as it resulted in better accuracy
- Experiment 4: CNN model from scratch
  - (64 x 64) image size
  - Accuracy on test set: ~50%
- Experiment 5: Pre-trained ResNet50
  - Accuracy on test set without data augmentation: ~85%
  - No data augmentation applied
  - No normalization applied
- **Final Model**: Pre-trained ResNet50
  - Accuracy on test set with data augmentation: ~85%
- Improve accuracy?
  - Ideal solution: More pictures
- Interesting trends...
  - Buildings with unique features had better classification



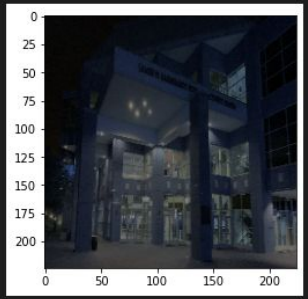
Sample outputs



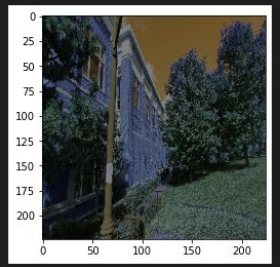
Prediction: 34 - Prospector
Image Path : ./val/34 - Prospector/Prospector (148).jpg



Prediction: 32 - Colvard
Image Path : ./val/32 - Colvard/Colvard (146).jpg



Prediction: 46 - SAC
Image Path : ./val/46 - SAC/SAC (318).jpg



Prediction: 12 - Smith
Image Path : ./val/69 - Student Union/Union (512).jpg

Sometimes it doesn't get it correct...

Conclusion

- Success!
- Understanding the AI “black box”
  - Unique buildings produce better results
  - Buildings often get confused with similar buildings (architecture)
- CNN are extremely powerful
  - Good performance on limited data
- What next?
  - More images
  - More diverse images
  - Better data collection approach
  - More buildings covered

Presentation sources

- P. Bezak, "Building recognition system based on deep learning," 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR), 2016, pp. 1-5, doi: 10.1109/ICAIPR.2016.7585230.
- https://machinelearningmastery.com/how-to-choose-loss-functions-when-training-deep-learning-neural-networks/
- https://towardsdatascience.com/cross-entropy-loss-function-f38c4ec8643e
- https://opt-ml.org/papers/2021/paper53.pdf
- https://towardsdatascience.com/deep-learning-optimizers-436171c9e23f
- More found in research report!
- https://towardsdatascience.com/understanding-and-coding-a-resnet-in-keras-446d7ff84d33

Q & A

Thank you for listening! We will now take questions