Project Report

Creating a Neural Network to Classify Indoor Corridors

By Moritz Zeumer

Overview

- Related Literature
  - End-To-End Navigation with Branch Turning Support Using Convolutional Neural Networks
  - Have I Reached the Intersection
- Tutorial Projects
  - Classify Movie Reviews
  - Classify Newswire Topics
  - Classify Dog & Cat Images
- IntersectNet
  - CG Dataset
  - Network Structure, Training, and Results
- Results and Future Work

Related Literature

End-To-End Navigation with Branch Turning Support Using Convolutional Neural Network

- Development of a navigation system based on end-to-end learning
  - Branch detection and behavior change to reach different destinations
- Frontal camera images input & Vector to next target introduced after convolutional layers
- Pure pursuit algorithm to follow planned trajectory

- Mobile robot capable to follow different trajectories on unknown routes
- Mobile robot learned to navigate locally only with current target direction

Related Literature

Have I Reached the Intersection:
A Deep Learning-based Approach for Intersection Detection from Monocular Cameras

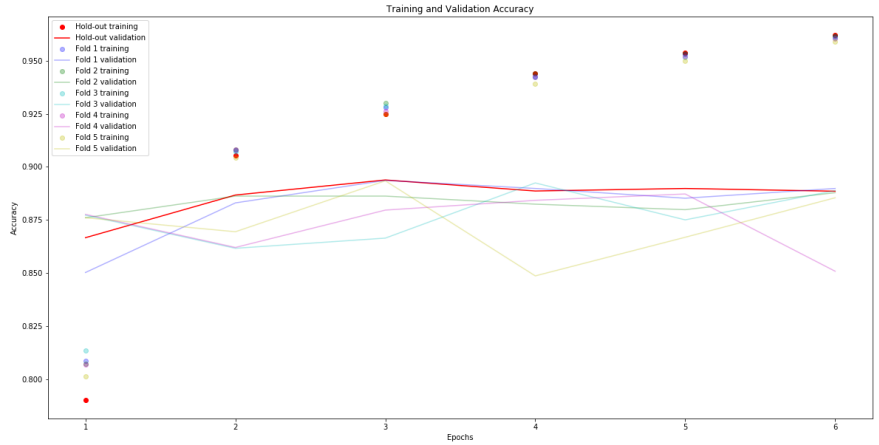
- Compound network of CNN and RNN: Long-Term Recurrent Convolutional Network
  - Convolutional Neural Network for special feature encoding
  - Recurrent Neural Network for temporal connection between sequence frames
- Trained on sequences from Oxford RobotCar & LARA traffic-light detection datasets

- Resulting network achieves ~92% accuracy and temporal connection leads to 2.5-5.5% improvement compared to single-frame networks

Tutorial Project – Classify Movie Reviews

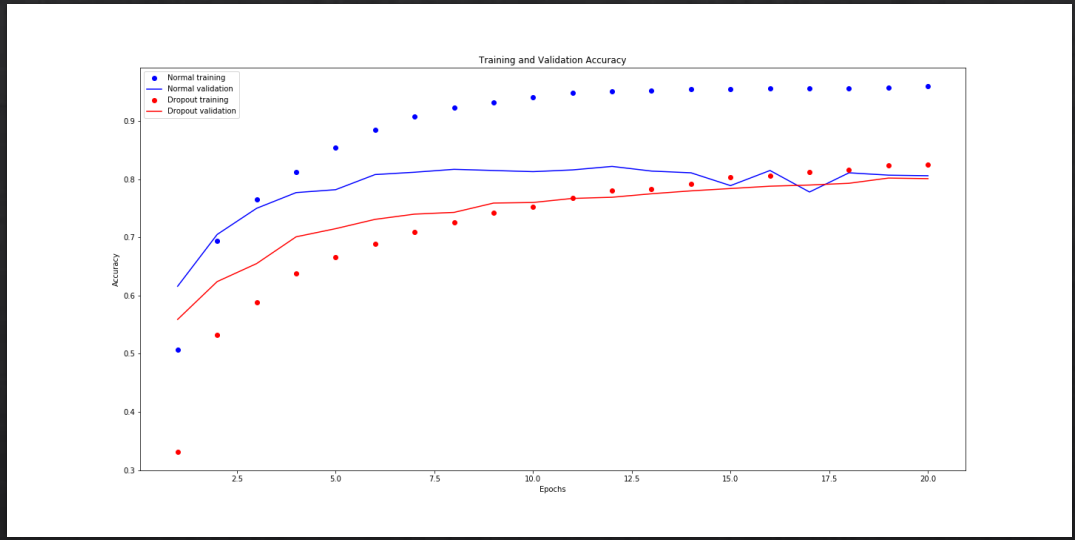
- Dataset of 50.000 positive & negative movie reviews (binary classification)
- Simple network with 3 fully connected layers (16x16x1) is trained using hold-out and k-fold cross-validation (6 epochs)
- Network Accuracy:
  - Hold-Out Model: 87,70%
  - K-Fold Model: 86,82%
- Models perform similarly. However, k-fold cross-validation is more robust with small datasets

Model	Split 1	Split 2	Split 3	Split 4	Split 5
1	**Validation**	*Training*	*Training*	*Training*	*Training*
2	*Training*	**Validation**	*Training*	*Training*	*Training*
3	*Training*	*Training*	**Validation**	*Training*	*Training*
4	*Training*	*Training*	*Training*	**Validation**	*Training*
5	*Training*	*Training*	*Training*	*Training*	**Validation**



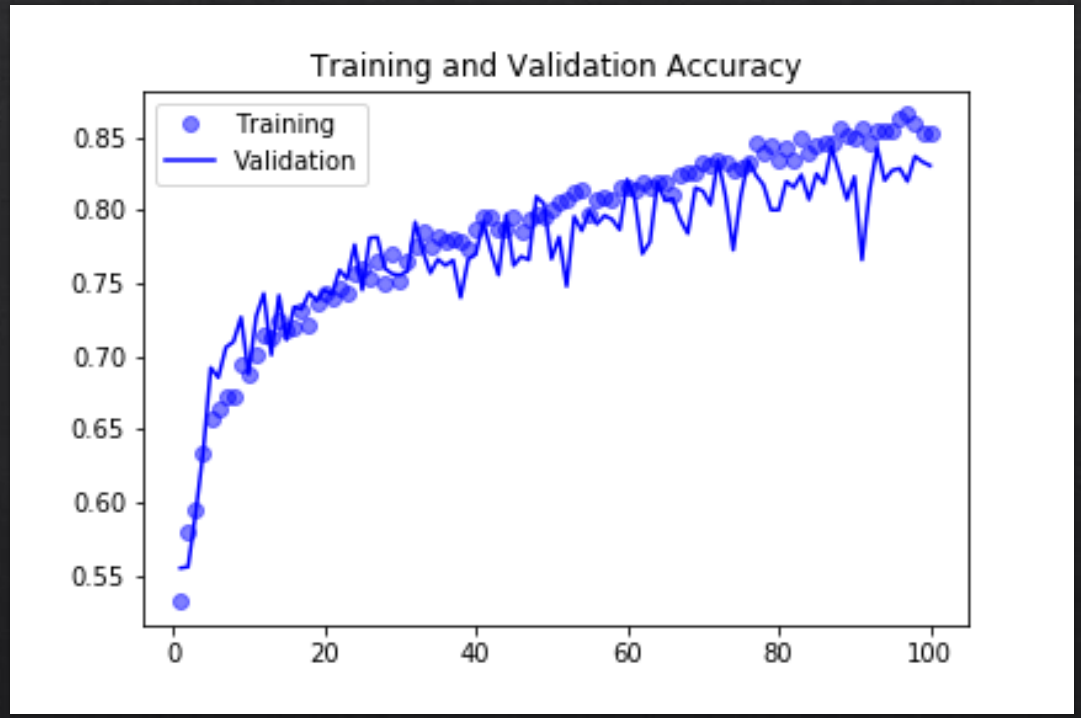
Tutorial Project – Classify Newswire Topics

- Dataset of 11.228 newswire articles in 46 topics (multiclass classification)
- Simple network with 3 fully connected layers (64x64x46) is trained with 50% dropout after each layer (20 epochs)
- Network Accuracy:
  - No-Dropout Model: 77,91%
  - Dropout Model: 77,07%
- Models perform similarly. However, the no-dropout model achieves peak performance faster, but is less resistant to overfitting



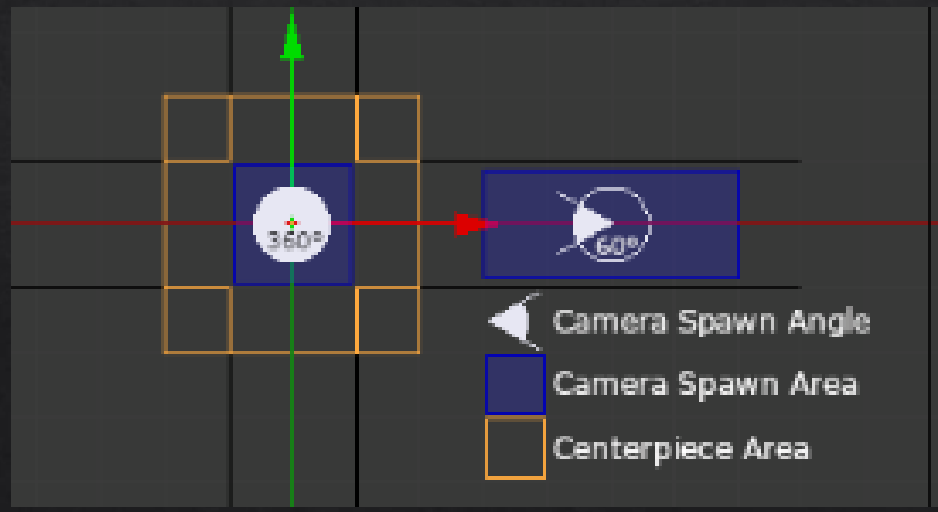
Tutorial Project – Classify Dog & Cat Images

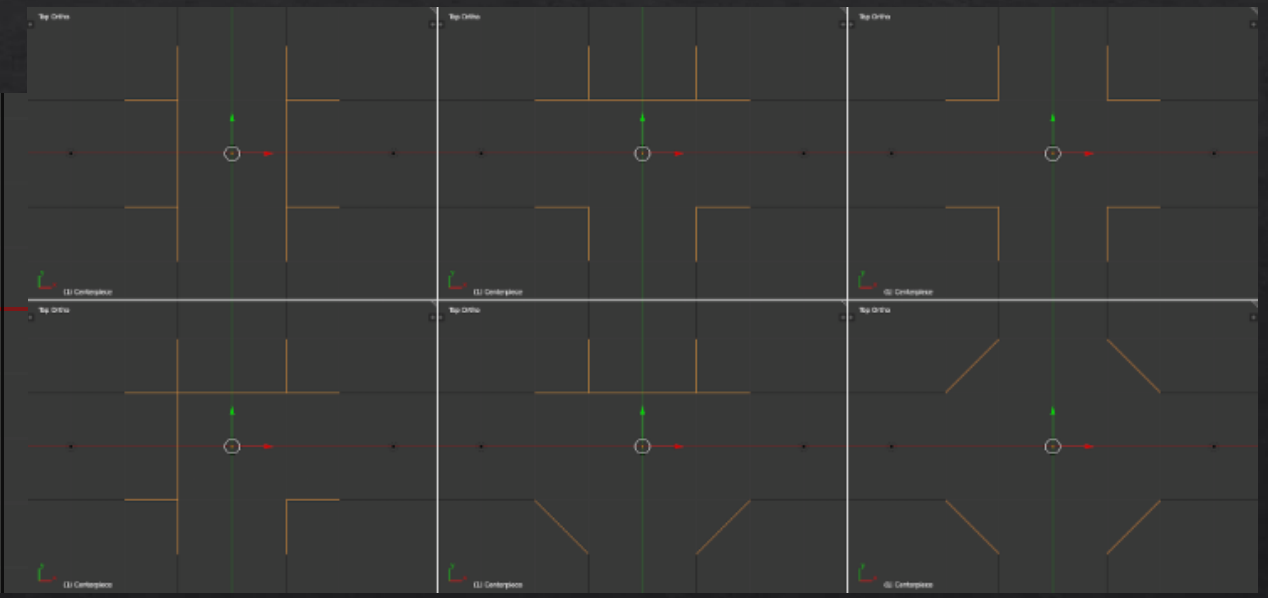
- Dataset of 4.000 images of dogs and cats (binary classification)
- Simple Convolutional network with 4 convolutional layers and 2 fully connected layers is trained with pooling layers after each convolutional layer and 50% dropout after the fully connected layer (100 epochs)
- Network Accuracy: 81%
- Model performs adequate considering the heavily reduced dataset (originally 25.000 images). Dataset preprocessing and augmentation through random permutation very beneficial for limited datasets



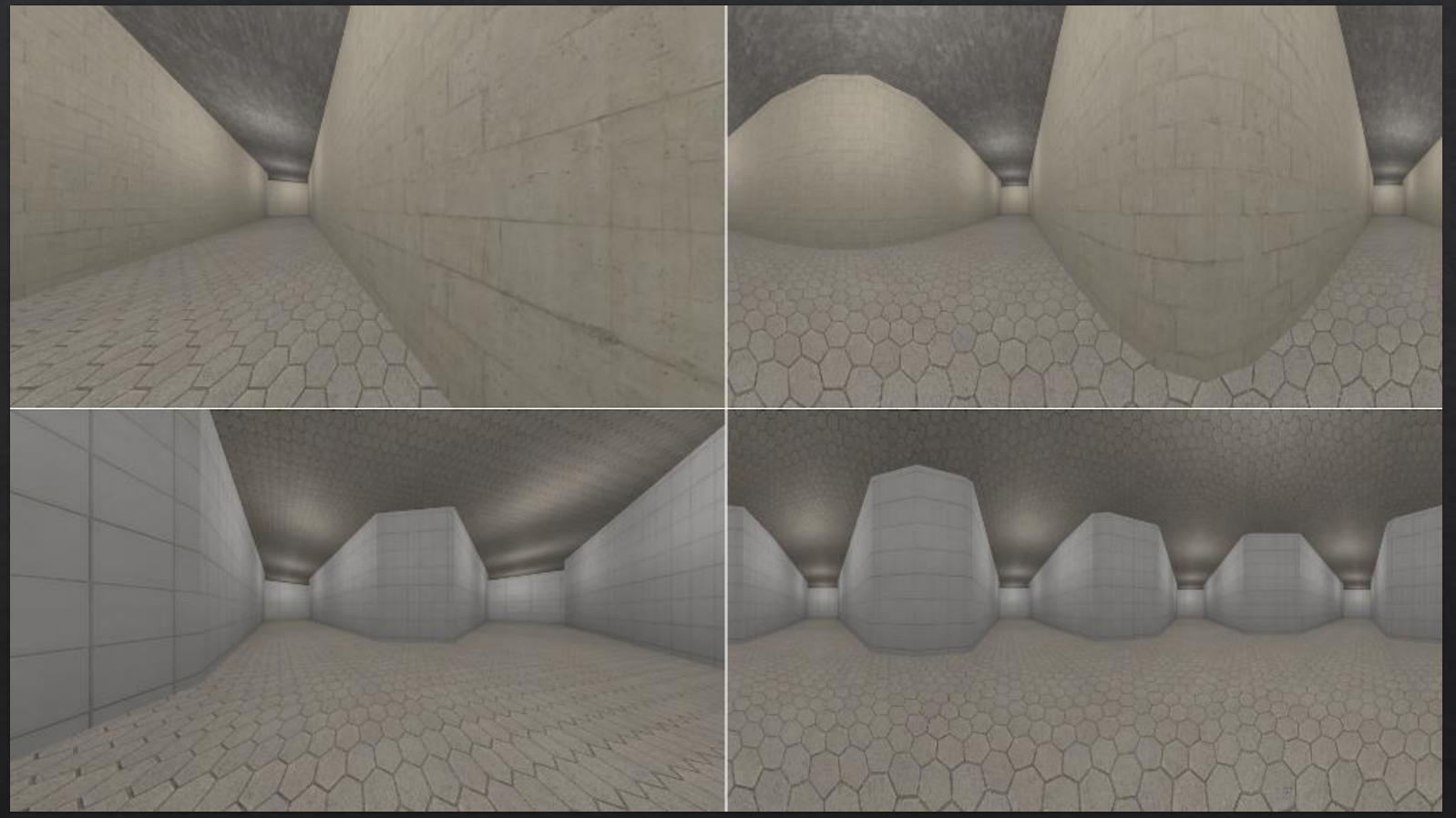
IntersectNet – CG Dataset

- No datasets available. Lack of time and resources to record real data => render CG Dataset
- Fixed layout with randomized centerpiece (random amount of connections rotated randomly)
- Camera created at semi-random location and rotation
- Layout textured with random textures



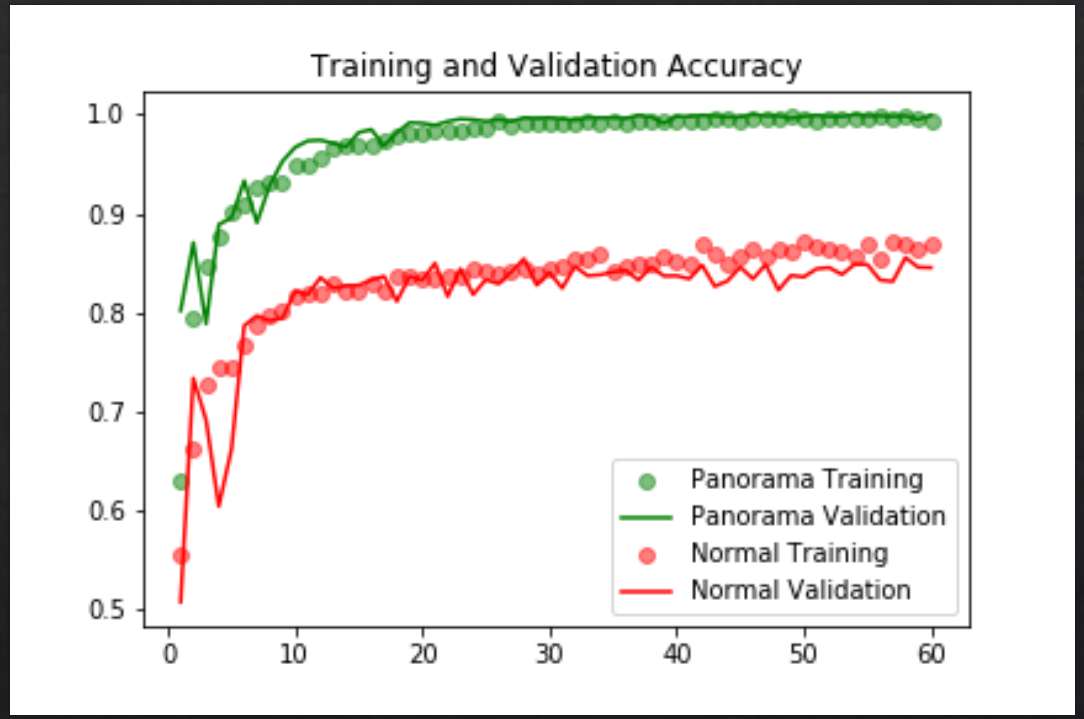


IntersectNet – CG Dataset



IntersectNet – Network Structure & Training

- Dataset of 9.000 images of corridors and intersections (binary classification)
- Simple Convolutional network with 4 convolutional layers and 2 fully connected layers is trained with pooling layers after each convolutional layer and 50% dropout after the fully connected layer (60 epochs)
- Accuracy Results:
  - Normal Images: 84.2%
  - Spherical Images: 99.9%
- The model trained with spherical images significantly outperforms the model trained with normal images.



IntersectNet – Results

- Model also tested with real images (12 corridors, 12 intersections, 6 other locations)
  - Only predicts 14/24 correct (58%)
  - Over-prediction of intersections
  - Possible reasons:
    - Not enough variety/quality in dataset
    - Wide open spaces/corridors (trained with 2m wide corridors)
    - Bright light sources and reflections
  - Fix: Improve CG image variety/quality or retrain with real image dataset

- Spherical imaging for autonomous navigation advantageous
  - Single sensor needed & no reliance on image sequences => faster detection/decision time

Thank You For Your Attention!