HUI WEI

(+1) 917 698 7063 ⋄ davidwei@nyu.edu

EDUCATION

Courant Institute of Mathematical Sciences, New York University May 2019
Master of Science in Computer Science

Beijing University of Posts and Telecommunications (BUPT) July 2017
Bachelor of Engineering in Telecommunication Engineering

RESEARCH EXPERIENCE

Multitask Prediction of Disease Onset from EHR data Jul.2019 - Present
Job title: Project Associate, Department of Population Health, NYU Langone Health

- Used 100 most common lab values of NYU EHR data to predict 283 diseases onset after one year.
- Grouped ICD-10 Code into 283 categories based on structural definition files, defined qualified patients for each disease, and reasonable labels from clinician diagnosis within outcome windows.
- Built two innovative CNNs. One captures temporal information of each lab value from 3 resolutions, then fuses temporal information for all lab values and resolutions. The other mixes information across different lab values at each temporal step using two vertical convolution layers, then captures temporal information using temporal convolution and pooling layers.
- Implemented LSTM model as a baseline.

Dementia Sub-type Prediction at the First Visit of Early Stages Feb.2019 - Present
Advisor: Prof. Narges Razavian, NYU Langone Health

- Designed and implemented Softmax Logistic Regression and Fully Connected Net for classifying four dementia sub-types at the first visit of early stages for demented patients, using demographic and biomarker data from NACC dataset as features and neuropathological results as labels.
- Proposed four Data Augmentation methods using clinician diagnosed labels, and utilized L1 Regularization to boost macro F1 score from 0.365 to 0.408, outperforming physicians.
- Analyzed reasons of clinician diagnosis errors and where the best model improves, using top features obtained from the weight matrix of each disease.

Track-Net May.2018 - Aug.2018
Advisor: Dr. Chenge Li, Prof. Yao Wang, NYU ECE Department

- Implemented and trained Detect-to-Track and Track-to Detect model on ImageNet dataset, achieving the accuracy stated in the original paper ($mAP = 82.3\%$).
- Fine-tuned Detect-to-Track and Track-to-Detect model on UADETRAC traffic video dataset.
- Proposed to use LSTM to enable the network to detect the dynamic bounding boxes for vehicles.

ID Card Detection and Information Extraction Sept.2015 - Feb.2016
Advisor: Prof. Bo Xiao, BUPT

- Used Gaussian filters to eliminate the noise in images to make it smoother
- Proposed to use Hough Line and Circle Transform to detect borders and corners of ID cards within different backgrounds, the accuracy is mAP = 63%.
- Enhanced the detection accuracy to mAP = 85%, using SIFT descriptors and matching algorithms between referred and query images.
- Extracted identity information contained in ID cards using OCR.

CLASS PROJECTS

CheXpertNet Spring 2019
Course: Deep Learning in Medicine

- Trained DenseNet on both frontal and lateral views from CheXpert dataset for 3 epochs to classify 14 diseases for each patient, achieving mean AUC score of 0.8236.
- Visualized the heatmap using Class Activation Map (CAM) to show which part is attended by the model when it makes a decision for each disease.
- Conducted three ablation studies and concluded: (1) lateral views help diagnosis for most diseases. (2) normalizing input images using statistics of specific dataset is better than statistics of ImageNet. (3) transfer learning benefits training.

Deep Segmentation for Mouse Embryo Brain Ventricles Fall 2018
Course: Machine Learning for Healthcare

- Trained VGG net as a classifier on a sliding window to detect whether it contains brain ventricles.
- Designed an ensemble V-net-like 3D CNN model to segment the whole ventricle in each sliding window, improving Dice based function from 0.7119 to 0.8956, compared with the graph-cut baseline.
- Leveraged spatially separated / cross constrained filter structures to reduce parameter of the model from 15M to 2M.

Learning to Generate Chairs with Convolutional Neural Networks Fall 2018
Course: Computer Vision

- Implemented an upside-down CNN model with conv layers and de-conv layers, to generate 2D CAD chair images and segmentation masks according to different chair type and view angles.
- The only implementation available is from the author in Caffe with Lua. Our implementation is cleaner and more up-to-date using Pytorch with Python.
- Found that 1) the network can create new views angles instead of remember those provided in the training dataset, 2) segmentation branch aids to fix artifact for the generation branch.

Transfer Learning for Text Classification Fall 2018
Course: Foundations of Machine Learning

- Utilized CNN model as a feature extractor and SVM as a classifier to solve binary and multi-class text classification problem on Amazon and Yelp review dataset.
- Used transfer learning to improve the robustness and effectiveness of the classification model, by training CNN on Amazon dataset, freezing it and training SVM on Yelp Binary and Multi-label dataset.

CycleGAN for Autonomous Driving Fall 2017
Course: Vision Meets Machine Learning

- Implemented CycleGAN model for driving pictures and videos to transfer styles between the game and the real world.

TECHNICAL SKILLS

Programming Languages	C++, Python, MATLAB, LaTeX, Markdown
Tools & Libraries	Pytorch, Pandas, Scikit-Learn, OpenCV, CVX