Computer Vision Nanodegree Syllabus

Master the computer vision skills behind advances in robotics and automation.

UDACITY
FOR ENTERPRISE

Table of Contents

Before You Start	3
Contact Info	3
Nanodegree Program Info	3
Project 1: Facial Keypoint Detection	4
Project 2: Automatic Image Captioning	5
Project 3: Landmark Detection and Tracking	6

Before You Start

Prerequisites: In order to succeed in this program, we recommend having significant experience with Python, and entry-level experience with probability and statistics, and deep learning architectures. Specifically, we expect you to be able to write a class in Python and to add comments to your code for others to read. Also, you should be familiar with the term "neural networks" and understand the differential math that drives backpropagation.

Educational Objectives: In this program, you'll learn the underlying math and programming concepts that drive pattern recognition, object and image classification tasks, and object tracking systems. This course will cover the latest in deep learning architectures used in industry, and you'll combine current computer vision and deep learning techniques to power a variety of applications. With the practical skills you gain in this course, you'll be able to program your own applications, extract information from any kind of image and spatial data, and solve real-world challenges.

Contact Info

While going through the program, if you have questions about anything you can email us at enterprise-support@udacity.com. For help from Udacity mentors and peers, please visit the Udacity classroom.

Nanodegree Program Info

Technical Requirements
Hardware Requirements: 64-bit operating system with 8GB of RAM, webcam, microphone
Software and Software Version Requirements: Text editor, Python 3.6, Jupyter Notebooks, Anaconda, PyTorch, Keras, OpenCV

LENGTH OF PROGRAM*: 4 months
FREQUENCY OF CLASSES: Self-paced

*This is a self-paced program and the length is an estimation of total hours the average student may take to complete all required coursework, including lecture and project time. Actual hours may vary.

Project 1: Facial Keypoint Detection

Use image processing techniques and deep learning techniques to detect faces in an image and find facial keypoints, such as the position of the eyes, nose, and mouth on a face. This project tests your knowledge of image processing and feature extraction techniques that allow you to programmatically represent different facial features. You'll also use your knowledge of deep learning techniques to program a convolutional neural network to recognize facial keypoints. Facial keypoints include points around the eyes, nose, and mouth on any face and are used in many applications, from facial tracking to emotion recognition.

Supporting Lesson Content: Introduction to Computer Vision

Lesson Title	Learning Outcomes
INTRODUCTION TO COMPUTER VISION	• Learn where computer vision techniques are used in industry • Prepare for the course ahead with a detailed topic overview. • Start programming your own applications!
IMAGE REPRESENTATION AND ANALYSIS	• See how images are represented numerically. • Implement image processing techniques like color and geometric transforms. • Program your own convolutional kernel for object edge-detection.
CONVOLUSIONAL NN LAYERS	• Learn about the layers of a deep convolutional neural network: convolutional, maxpooling, and fully-connected layers. • Build an CNN-based image classifier in PyTorch. • Learn about layer activation and feature visualization techniques.
FEATURES AND OBJECT RECOGNITION	• Learn why distinguishing features are important in pattern and object recognition tasks. • Write code to extract information about an object's color and shape. • Use features to identify areas on a face and to recognize the shape of a car or pedestrian on a road.

Supporting Lesson Content: Introduction to Computer Vision (Continued)

Lesson Title	Learning Outcomes
FEATURES AND OBJECT RECOGNITION	• Implement k-means clustering to break an image up into parts. • Find the contours and edges of multiple objects in an image. • Learn about background subtraction for video.

Project 2: Automatic Image Captioning

Combine CNN and RNN knowledge to build a deep learning model that produces captions given an input image. Image captioning requires that you create a complex deep learning model with two components: a CNN that transforms an input image into a set of features, and an RNN that turns those features into rich, descriptive language. In this project, you will implement these cutting-edge deep learning architectures.

Supporting Lesson Content: Advanced Computer Vision and Deep Learning

Lesson Title	Learning Outcomes
ADVANCED CNN ARCHITECTURES	• Learn about advances in CNN architectures. • See how region-based CNN's, like Faster R-CNN, have allowed for fast, localized object recognition in images. • Work with a YOLO/single shot object detection system.
RECURRENT NEURAL NETWORKS	• Learn how recurrent neural networks learn from ordered sequences of data. • Implement an RNN for sequential text generation. • Explore how memory can be incorporated into a deep learning model. • Understand where RNN's are used in deep learning applications.

Supporting Lesson Content: Advanced Computer Vision and Deep Learning (Continued)

Lesson Title	Learning Outcomes
ATTENTION MECHANISMS	• Learn how attention allows models to focus on a specific piece of input data. • Understand where attention is useful in natural language and computer vision applications.
IMAGE CAPTIONING	• Learn how to combine CNNs and RNNs to build a complete captioning model. • Implement an LSTM for caption generation. • Train a model to predict captions and understand a visual scene.

Project 3: Landmark Detection and Tracking

Use feature detection and keypoint descriptors to build a map of the environment with SLAM (simultaneous localization and mapping). Implement a robust method for tracking an object over time, using elements of probability, motion models, and linear algebra. This project tests your knowledge of localization techniques that are widely used in autonomous vehicle navigation.

Supporting Lesson Content: Object Tracking and Localization

Lesson Title	Learning Outcomes
OBJECT MOTION AND TRACKING	• Learn how to programmatically track a single point over time. • Understand motion models that define object movement over time. • Learn how to analyze videos as sequences of individual image frames.

Supporting Lesson Content: Object Tracking and Localization (Continued)

Lesson Title	Learning Outcomes
OPTICAL FLOW AND FEATURE MAPPING	• Implement a method for tracking a set of unique features over time. • Learn how to match features from one image frame to another. • Track a moving car using optical flow.
GRAPH SLAM	• Identify landmarks and build up a map of an environment. • Learn how to simultaneously localize an autonomous vehicle and create a map of landmarks. • Implement move and sense functions for a robotic vehicle.