# DSC 180 A01 - EXPLAINABLE AI DISCUSSION 7: GRAD-CAM FOR XAI

Jürgen P. Schulze, Ph.D.
University of California, San Diego
Fall Quarter 2020

# ANNOUNCEMENTS

- Instructor office hour Fridays 10-11am
- Replication Report due Dec 11$^{th}$ at 11:59pm
- Next Participation due Dec 1$^{st}$ at noon

- Today: Guest speaker **Kamran Alipour** presenting on the Explainable AI (XAI) project

# DONE: TASKS THIS WEEK

**Reading**

- Read chapter 5.4 of Deep Learning with Python (by François Chollet).

- Skim chapters 5.1-5.3 as needed to be able to follow chapter 5.4.

**Participation Assignment (due November 24th at noon)**

Book chapter 5.4 describes three ways for visualizing what neural networks learn: intermediate activations, CNN filters, heatmaps. Summarize each of them in your own words in 2-3 sentences each.

# TO DO: TASKS FOR NEXT WEEK

**Reading**

- Read the original Grad-CAM paper "Grad-CAM: Why did you say that? Visual Explanations from Deep Networks via Gradient-based Localization"
  - https://arxiv.org/pdf/1610.02391v1.pdf

**Participation**

Answer the following questions (2-3 sentences each):

- What is the goal of the Grad-CAM algorithm?

- How does the Grad-CAM algorithm work?

- How does Grad-CAM differ from Class Activation Mapping (CAM)?