**机器视觉**

**课程名称（中文）：计算机视觉与多媒体信息处理**

**课程名称（英文）：Computer Vision and Multimedia Information Processing**

**课程性质：专业选修**

**学分：3**

**学时：72，其中理论学时：36，实践（实验）学时：36**

**授课对象：本科生**

**先修课程：高等代数，线性代数，概率论**

**授课语言：中文**

**开课院系：数据科学与工程学院**

**课程网址：无**

**撰写人：陆雪松**

**审定人：钱卫宁**

**一、课程简介（中文）**

计算机视觉核心是使用信息技术来模拟人类视觉的功能。它的研究内容非常广泛，包括图像识别，目标检测，语义分割，实例分割，图像生成等。近年来该领域研究热点是利用深度学习技术来解决各类计算机视觉的任务，这也是本课程讲授的主要内容。此外，计算机视觉技术可以与其他多媒体数据相结合，产生一些更加有趣的任务，包括图片描述、视频处理等。通过本课程的学习，有助于了解计算机视觉及相关多媒体技术的发展前沿，掌握使用深度学习技术解决计算机视觉任务的基本能力。

1. **课程目标**

本课程分为理论学习和实践两部分，通过对该课程的学习，使学生掌握使用深度学习技术处理计算机视觉及相关多媒体信息任务的基本方法，培养学生开发和训练深度学习模型的能力。

**三、教学内容、学时分配和作业要求**

（一） 计算机视觉简介（2）

主要内容：计算机视觉的基本任务和相关技术的发展历史

（二） 图像分类（2）

主要内容：图像分类任务的定义，KNN和线性分类器

（三） 损失函数和优化（2）

主要内容：SVM和Softmax分类器，正则化，优化

（四） 神经网络和反向传播（4）

主要内容：前馈神经网络的基本结构，计算图，反向传播和链式规则

（五） 卷积神经网络（4）

主要内容：卷积神经网络的基本结构，各层的反向传播，理解卷积神经网络

（六） 神经网络的训练（4）

主要内容：激活函数，数据预处理，参数初始化，Batch Normalization及其演化，优化方法，学习率，正则化，超参数的选择

（七） 深度学习硬件和软件（2）

主要内容：CPU、GPU、TPU的基本概念，Pytorch，TensorFlow

（八） 卷积神经网络结构的演化（2）

主要内容：LeNet，AlexNet，VGG，Inception，ResNet，迁移学习

（九） 图像风格迁移（2）

主要内容：图片纹理，Deep Dream

（十） 循环神经网络和图像描述（4）

主要内容：RNN的结构，图像描述，梯度爆炸和梯度消失

（十一）目标检测和图像分割（2）

主要内容：目标检测，语义分割，实例分割

（十二）生成模型（4）

主要内容：PixelRNN/CNN，VAE，GAN

（十三）自监督学习（2）

主要内容：自监督学习基本概念，Pretext Tasks，对比学习

教学要求：掌握开发深度学习模型的基本能力，掌握使用深度学习技术解决计算机视觉任务的基本能力

重点、难点：对计算机视觉及其算法产生兴趣，具备基本的深度学习模型开发和训练能力。

其它教学环节：课内和课外上机实践。

作业要求：独立完成。

**四、教材、参考书目或其他学习材料**

**参考书目：**

Goodfellow I, Bengio Y, Courville A. Deep learning. MIT press. 2016.

**五、考核办法与评价结构比例**

1．平时成绩：50%

2．期中考试：20%（闭卷）

3. 期末考察：30%