

# VISION@OUC 第一阶段任务（入学前）

郑海永/俞智斌/王楠

<http://vision.ouc.edu.cn>

2017 年春

## 目录

1	基础能力	1
1.1	技术	1
1.2	专业	2
1.3	要求	2
1.4	考核	3
2	研究方向	3
2.1	郑海永	3
2.2	俞智斌	5
2.3	王楠	7

## 1. 基础能力

### 1.1 技术

现阶段学习不是为了考试而是实用！所以，看懂无用，会用才好！

**操作系统** GNU/Linux，推荐 Ubuntu。

**编程语言** C/C++ 和 Python，此项非常重要！

关于程序设计，强烈推荐北京大学专项课程“[程序设计与算法](#)”（[资源下载](#)）

关于编程测试，基本要求：半年内 LeetCode (<https://leetcode.com>) 做两遍（第一遍可以参考别人代码，第二遍务必自己独立完成）。

**代码管理** GitHub，推荐 [GitHub Guides](#)。

**文档排版** L<sup>A</sup>T<sub>E</sub>X，推荐 [lshort -A short introduction to L<sup>A</sup>T<sub>E</sub>X2e](#)。

## 1.2 专业

现阶段学习不是为了考试而是实用！所以，看懂无用，实践才对！

**Machine Learning** 建议：《机器学习》，周志华，清华大学出版社，2016 年。

强烈建议认真学习 Andrew Ng 在 Coursera 上的“[Machine Learning](#)”课程，网上有很多相关的学习笔记供参考，学习时须加强**理论推导**和**实验分析**。“机器学习”是人工智能的基石，作为重中之重，请重点理解并加强实践！

**Deep Neural Networks** 建议：《深度学习》，Ian Goodfellow, Yoshua Bengio, Aaron Courville。

强烈建议认真学习 Fei-Fei Li 的“[CS231n: Convolutional Neural Networks for Visual Recognition](#)”课程，网上还有关于这门课程的学习讨论和2016 视频。“深度学习”是当前人工智能最为实用和有效的方法，很多工作都基于此，建议重点学习并实践实战！

**数字图像处理** 建议：Digital Image Processing (3rd edition), by Rafael Gonzalez and Richard Woods, Prentice Hall, 2008.

**计算机视觉** 建议：Computer Vision: Algorithms and Applications, by Richard Szeliski, Springer, 2010.

## 1.3 要求

1. 每人创建 GitHub 帐号进行学习过程中的代码和文档（包括笔记等）管理。

2. 所有的笔记等文档全部使用  $\text{LATEX}$  来撰写。
3. 所有的专业知识学习要求必须有对应的代码和实验及记录等文档。

## 1.4 考核

1. **GitHub**：入学时我们会详细查阅每个人的 GitHub 中的代码和文档等来评估是否对学习过程进行了详细的记录和整理，尤其自己的思考和问题等。
2. **答辩**：入学后一周内进行现场答辩，将自己对上述所学及下述感兴趣研究方向的方方面面进行汇报，并接受老师质疑，时间每人不超过 1 小时。

## 2. 研究方向

### 2.1 郑海永

主要的研究方向为**图像识别**或**图像分类**，在此问题下又分为以下两个方面：

1. **中文**：细粒度图像分类/图像精细识别

**English**: Fine-Grained Visual Classification / Fine-Grained Image Classification / Fine-Grained Image Recognition

**简要说明**：通过图像识别是狗还是鸟相对比较容易，而如果从鸟中识别是哪一种鸟（生物分类阶元：界门纲目科属种），则难上加难（可能只有生物分类专家才能做到），如图1所示。

**主要方法**：深度神经网络（Deep Neural Network）

2. **中文**：不平衡数据（图像）分类

**English**: Imbalanced Data (Image) Classification

**简要说明**：类别不平衡现象广泛存在于现实生活中的大量实际应用中，比如你看到“老虎”的样本数要远少于你看到“狗”的样本数，那么对于多数类样本和少数类样本的分类问题，就很难用传统的机器学习方法来解决。例如，



图 1: 图像精细识别

假设一个数据集包括 2 类，其中多数类样本数为 99、少数类样本数为 1，那么分类器可以直接忽视掉少数类（全部样本识别为多数类）便可以得到 99% 的识别准确率，尽管这个识别准确率已经算很高了，但是对于少数类的识别准确率却为 0，而有时候少数类的识别准确率可能更重要，比如诊断癌症、信用诈骗等。

**主要方法：**类别不平衡学习（Class-Imbalance Learning）、生成式对抗网络（Generative Adversarial Network）

### 3. 中文：图像质量评价

**English:** Image Quality Assessment (IQA)

**简要说明：**图像质量评价是图像处理中的基本技术之一，主要通过对图像进



行特性分析研究，然后评估出图像优劣（图像失真程度）。图像质量评价在图像处理系统中，对于算法分析比较、系统性能评估等方面有着重要的作用。近年来，随着对数字图像领域的广泛研究，图像质量评价的研究也越来越受到研究者的关注，提出并完善了许多图像质量评价的指标和方法。

**主要方法：**主观评价和客观评价，客观评价包括全参考 (Full-Reference, FR)、部分参考 (Reduced-Reference, RR) 和无参考 (No-Reference, NR) 三种类型。

## 2.2 俞智斌

### 1. 中文：深度图像估计/图像深度估计。

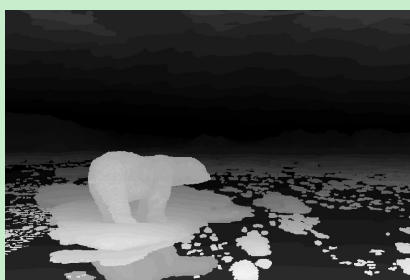
**English:** Depth Map Prediction / Depth Map Estimation / Depth Prediction / Depth Estimation.

**简要说明：**通过一副或多幅图像预测/估计图像中每个物体（像素）的深度信息。如图2所示，图2(a)表示原始图像，图2(b)为真实的深度图（可以通过微软 Kinect 等设备采集获得），图2(c)为通过所研究的方法（算法）由图2(a)所得到的预测深度图。实际应用时，我们只通过原始图像就可以计算得到（预测）深度图，那么方法（算法）的好坏就由预测/估计的准确性来评判。

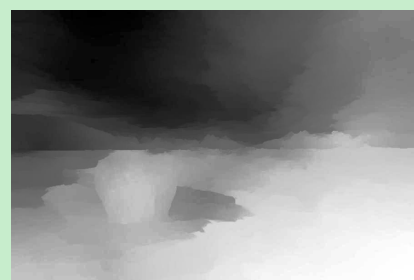
**主要方法：**深度神经网络 (Deep Neural Network)



(a) 原图



(b) 真实深度图



(c) 预测深度图

图 2: 深度图像估计/图像深度估计

### 2. 中文：图像复原。

**English:** Image Restoration.

**简要说明：**图像复原就是利用退化过程的先验知识，去恢复已被退化图像的本来面目。成像系统受各种因素的影响，导致了图像质量的降低，称之为图像退化。因素包括：传感器噪声、摄像机聚焦不佳、物体与摄像机之间的相对移动、随机大气湍流、光学系统的象差、成像光源和射线的散射等。退化基本表现：图像模糊。如图3所示为北京灰霾的照片及其去雾结果，图4所示为模糊图像及其去模糊结果和去模糊所用到的模糊核。图像去雾（dehaze）和图像去模糊（deblur）可以应用图像复原方法来处理。



图 3: 图像去雾 (Dehaze)



图 4: 图像去模糊 (Deblur)

### 3. 中文：基于图像的参数回归。

English: Paramter Regression.

**简要说明：**图像中蕴含了成像介质的很多信息（参数），如何从图像中获取到这些信息是本研究方向的核心问题。此方向的研究思路比较新颖，并不能查到太多资料，如有问题请单独联系俞老师。

**主要方法：**深度神经网络（Deep Neural Network）

## 2.3 王楠

1. 嵌入式编程及电路设计（机器人），此方向“软硬结合”，需要有单片机或电路设计经验。

**技能要求：**1) 掌握 C/C++ 编程，会进行图形化界面编程。2) 对嵌入式系统有浓厚兴趣爱好，有 ARM 或 FPGA 开发经验者优先最佳。3) 使用 Altium Designer（DXP、Protel 99）等电路设计软件制作电路板。

在这里，你要做的事，**制作自己的机器人……当然，越智能越好。**

2. 数字图像处理，主要侧重模糊图像中目标边缘检测等。

**技能要求：**1) 数学基础扎实。2) 对“信号与系统”、“数字信号处理”有较好的领悟。3) 掌握 C/C++ 编程。

在这里，你要做的事，**结合机器人研发，探索如何在模糊不清的环境中进行目标检测，并进一步帮助机器人完成基于图像的路径规划等问题。**