

# COMP9517: 计算机视觉

## 2023年第二学期

### 小组项目规范 可获得的最高分

**数： 40**

小组项目的**分数占课程总分的40%**。

项目工作在第6-10周进行，演示和报告在第10周到期。

有关评分的详细信息，请参考单独的评分标准。

提交说明和演示时间表将在以后发布。

#### 简介

小组项目的目标是在一个由4-5名学生组成的团队中与同伴合作，解决一个计算机视觉问题，并以口头和书面形式展示解决方案。

在第6-9周，每个小组可以在每周五下午2-3点的常规咨询会议上与他们指定的导师会面一次，讨论进展情况并获得反馈。

小组项目要由每个小组单独完成。不要复制其他小组的想法或任何材料。如果你在某个任务中使用了公开的方法或软件，这些方法或软件必须有适当的出处/参考。不这样做就是抄袭，将根据课程大纲中描述的新南威尔士大学规则进行处罚。

请注意，我们只给那些为了这个项目的目标而开发新东西或尝试更多以前没有使用过的最先进的方法的小组打高分。我们并不期望你从头开始开发，但你越是使用或建立在现有的代码上（会被检查），分数就越低。我们确实希望你能表现出创造性，并在你在课程中或从计算机视觉文献中学到的想法的基础上进行开发。

---

#### 描述

两个重要且具有挑战性的计算机视觉任务是现实世界图像或视频中的物体检测和分类。应用实例包括监控、交通监测、机器人、医疗诊断和生物学。

在许多应用中，大量和复杂的数据使得人类不可能对相关的图像信息进行准确、完整、高效和可重复的识别和分析，因此需要完全自动化。

这个小组项目的目标是开发和评估用于检测和评估的方法。

对野生动物图像中的动物进行分类。具体来说，在这个项目中，我们将关注两类动物：企鹅和海龟。我们面临的挑战是开发能够准确和有效地分析图像的方法。

## 任务

### 数据集

该小组项目使用的数据集是Kaggle提供的企鹅与海龟的数据集（见本文末尾的参考资料）。它包括一个由500张图片组成的训练集和一个由72张图片组成的验证集。每张图片都包含一只企鹅或一只海龟，在任意位置，如相应的注释文件中所示。

### 探测

第一个任务是检测和定位每张图片中的动物。具体来说，任务是开发一种方法，可以将数据集中的任何图像作为输入，并产生一个边界框作为输出（x\_min、y\_min、宽度、高度，均为像素）。

这取决于你是把它作为一个独立的任务来解决，还是先解决分类任务（接下来描述），然后用预测的类别标签来通知检测（因为这允许为每个类别采用一个更专门的检测器），甚至是你是否以某种方式联合解决这两个任务。

### 分类

第二项任务是对每张图片中的动物进行分类。具体来说，这项任务是开发一种方法，可以将数据集中的任何图像作为输入，并产生一个类别标签作为输出（1 = 企鹅，2 = 乌龟）。

这取决于你是把它作为一个独立的任务来解决，还是先解决检测任务（如上所述），然后用预测的边界框来通知分类（因为这样可以使注意力集中在动物身上而忽略更大的背景），甚至你是否以某种方式把这两个任务联合起来解决。

### 方法

许多传统和/或基于机器/深度学习的计算机视觉方法可用于这些任务。你面临的挑战是使用课程中教授的概念和文献中的其他方法来开发自己的方法并评估其性能。

一些流行的检测和分类方法的代码是公开可用的。你可以研究它们以获得灵感，但你不应该直接使用它们（我们会检查你是否使用了现有的代码，见上面和下面的说明）

。

虽然我们希望你从头开始开发所有的东西，但我们确实希望看到一些新的方法组合，或现有方法的一些调整，或使用更多以前没有尝试过的最先进的方法来解决特定的问题。

由于这里几乎有无限多的可能性，所以不可能给出详细的标准，但作为一般的指导原则，你越是自己开发，而不是直接复制自

在其他地方，越好。在任何情况下，都要引用你的来源。

### 培训

如果你的方法需要训练（也就是说，如果你使用有监督的而不是无监督的检测和分类方法），你可以使用训练集（500张图片）来达到这个目的。即使你的方法不需要训练，它们可能有超参数，你需要微调以获得最佳性能。在这种情况下，你也必须使用训练集，而不是验证集，因为在训练/微调和测试中使用（部分）相同的数据会导致有偏见的结果，不能代表实际性能。

### 测试

对于你的方法的测试，你必须使用验证集（72张图片）。为了评估该方法的整体性能，计算并报告以下指标。

检测性能：对于每张验证图像，计算预测边界框的中心位置与相应的真实边界框的中心位置之间的距离（可从注释文件中获得），并报告所有验证图像上的距离的平均值和标准偏差。同时计算每个验证图像的预测边界框和其相应的真实边界框的交集（IoU），并报告其平均值和标准偏差。

分类性能：对于每个验证图像，使用真实的类标签（可从注释文件中获得）来确定预测的类标签是否正确，并报告分类结果的混淆矩阵。由此，计算并报告你的方法的准确率、精确度、召回率和F1分数。

在你的演示和书面报告中展示这些量化分数（见下面的交付物），同时展示成功检测和分类的代表性例子，以及你的方法失败的例子（一般来说，没有任何方法能产生100%的完美结果）。请解释一下为什么你认为你的方法在这些情况下失败了。

### 视觉化

除了定量检测（如上所述），你的方法还必须显示检测和分类结果。也就是说，对于每张图片，它不仅应该检测和分类动物，还应该在图片上画出其相应的边界框和类别标签。

---

## **可交付的成果**

小组项目的成果是：1) 一个小组视频演示；2) 一个小组报告。两者都在第10周到期。

关于这两项成果的更多详细信息：

### 视频演示

每个小组将准备一个最多 10 分钟的视频演示，展示他们的工作。演示必须从介绍问题开始，然后解释所使用的方法。

方法，展示所获得的结果，并讨论这些结果以及对未来改进的想法。这一部分的介绍应该采用简短的PowerPoint幻灯片的形式。在这一部分之后，演讲应该包括方法/软件的实际演示。当然，有些方法可能需要很长的时间来计算，所以你可以录制一个现场演示，然后进行编辑，使其不超过时间。

整个演示必须以视频的形式（720p或1080p mp4格式），最多 10分钟（超过的部分将被砍掉）。所有小组成员都必须展示（如果不是这样，可能会被扣分），但由你决定谁展示哪一部分（介绍、方法、结果、讨论、演示）。为了让我们核实所有小组成员确实都在展示，每个展示自己部分的学生必须在展示的一角出现（现场录音，而不是静态的头像），而且当他们开始展示时，必须提到自己的名字。

使用PowerPoint本身的视频录制功能（例如[见本教程](#)）或使用其他录制软件，如[OBS Studio](#)、[Camtasia](#)、[Adobe Premiere](#)和其他许多软件，可以很容易地实现叠加网络摄像头的录制。这取决于你（取决于你的偏好和经验）使用哪种软件，只要最终的视频满足上述要求就可以了。

还要注意的，视频文件可以很容易地相当大（取决于使用的压缩级别）。为了避免本课程的存储问题，每组的视频上传限制为100MB，这对于这种类型的演示来说应该是绰绰有余。如果你的视频文件更大，可以使用[HandBrake](#)等工具以更高的压缩率重新编码。

在第10周的预定上课时间，即2023年7月31日星期一中午12-2点和/或2023年8月4日星期五下午1-3点，视频演示将展示给导师和讲师，导师和讲师将对其进行标记，并向小组成员提出相关问题。其他学生也可以收看并提出问题。因此，每个小组的所有成员都必须在其视频播放时在场。在接近第10周时，将制作并公布一份名册，显示每个小组的展示时间。

## 报告和代码

每个小组还将在2023年8月4日18:00 AEST之前提交一份报告（[2栏式的IEEE格式](#)，最多10页文本和任意数量的参考文献）以及源代码。在2023年8月4日18:00:00 AEST之前，每组还将提交一份报告（文本不超过10页，以及任意数量的参考文献），并附上源代码。

报告必须以PDF文件形式提交，并包括：

1. 介绍：讨论你对任务规范和数据集的理解。
2. 文献回顾：回顾文献中的相关技术，以及任何必要的背景，以了解你选择的方法。
3. 方法：激励并解释你所实施的方法的选择，必要时使用相关的参考资料和理论。
4. 实验结果：解释你用来评估所开发的方法的性能的实验装置和你获得的结果。



5. 讨论：提供对结果和方法性能的讨论，特别是方法失败的原因（如果适用）。
6. 结论：总结成功/不成功的地方，并对未来的工作提出建议。
7. 参考文献：列出你工作中使用的文献参考和其他资源。项目中使用的**所有**外部资源（包括网站）都必须有参考文献。参考文献部分不计入10页的限制。

所开发软件的完整源代码必须以ZIP文件的形式提交，并与报告一起由评委进行评估。因此，提交的文件必须包括所有必要的模块/信息以方便运行代码。难以运行的软件或不能产生所展示的结果，将导致扣分。源代码（ZIP）加报告（PDF）的上传限制为100MB。请注意，这个上传限制与视频上传限制是分开的（每个都是100MB）。

抄袭检测软件将被用来筛选所有提交的材料（报告和源代码）。不仅要在提交的材料之间进行配对比较，还要与往年的类似作业（如果适用）和公开的材料进行比较。新南威尔士大学的抄袭政策请参见课程大纲。

### 学生捐款

作为一个小组，你们可以自由地在小组成员之间进行分工，但所有小组成员必须在方法开发、编码、制作视频和撰写报告方面做出大致相同的贡献。例如，如果一些小组成员只准备视频和报告而不对方法和代码做出贡献，那是不可接受的。

学期结束时，将举行一次在线调查，让学生匿名评估其小组成员对项目的相对贡献。调查结果将只报告给LIC和课程管理员，如果有足够的证据表明个别学生的贡献大大低于其他组员，他们可以酌情调整项目的最终分数。

---

### **参考文献**

Kaggle.Penguins Versus Turtles：用边界框进行图像分类.2023.  
<https://www.kaggle.com/datasets/abbymorgan/penguins-vs-turtles>

---

**版权所有**：新南威尔士大学CSE COMP9517团队。复制、出版、张贴、分发或翻译本作业是对版权的侵犯，并将被提交给新南威尔士大学学生行为与诚信处处理。

**发布**：2023年6月30日