

第二章 计算机视觉

目标：

- 1. 掌握计算机视觉的基本概念；
- 2. 知道计算机视觉的典型应用；
- 3. 可以熟练掌握计算机视觉的相关应用；
- 4. 能够利用人工智能开源框架开发计算机视觉相关应用；
- 5. 建立起利用计算机视觉解决生活中所碰到的问题的意识；
- 6. 在日常学习生活中，建立起保护隐私信息（人脸、指纹、虹膜等）的意识，养成主动保护自己及他人的个人隐私数据的习惯。

在众多的科幻题材的电影中，都少不了机器人的角色，机器人有个很重要的能力，就是像人一样能看到这个世界，并且能够看懂这个世界，机器人甚至能够比人看到的更多，能够实时联网对所看到的对象进行检索和分析。

计算机视觉是人工智能感知世界的主要手段之一。相机的发明，让人类能够记录美好的世界，数码相机的发明，让人们能够数字化这个视觉世界，随着计算机视觉的发展，必将会实现数字世界到现实世界的融合，人工智能能够和人一样通过视觉感知世界，理解世界。那么到底什么是计算机视觉？计算机视觉有哪些应用？让我们一起开始探索吧！

一、计算机视觉初体验

表 2.1.1 计算机视觉之光学文字识别体验备选工具

环境	工具名称	功能简介	获取方式
Windows 操作系统	PandaOCR	支持各种场景下的文字识别，如图片识别、截图识别、剪切识别，支持多个主流识别引擎，支持翻译。	https://github.com/miaomiaosoft/PandaOCR/releases
移动终端下（IOS 和 Android）	讯飞输入法	讯飞输入法除了强大的语音输入功能外，还有拍照输入的功能，可以实现文字、翻译、公式、文档、表格、作业批改等功能	相应的市场输入讯飞输入法下载。或访问网址： https://srf.xunfei.cn/

计算机视觉体验案例 2-1：光学文字识别（OCR）

在体验时，可以根据信息环境的不同，选择最适合自己的环境进行体验，在此，提供 Windows 操作系统环境和移动终端环境，备选体验工具如表 2.1.1 所示。

1. PandaOCR 体验步骤

- （1）在 Windows 操作系统下访问 PandaOCR 的网址，如表 2.1.1 所示；
- （2）下载 PandaOCR，并解压缩；
- （3）双击打开 PandaOCR，免安装，解压后直接运行使用；
- （4）体验图片识别、截图识别或剪切识别功能；
- （5）根据体验填写表 2.1.2



图 2.1.1 PandaOCR 界面

2. 讯飞输入法的拍照输入体验步骤

- （1）在移动终端上，如苹果手机、ipad、Android 手机、平板电脑，访问相应的 APP 市场，并下载讯飞输入法，以下步骤以 IOS 下的操作为例，其他系统略有不同；
- （2）在讯飞输入法，“我的”，“工具箱”中启用拍照输入功能；
- （3）在讯飞输入法，“我的”，“工具箱”中“拍照输入”点击“开始使用”；
- （4）填写表 2.1.2 的体验记录表。

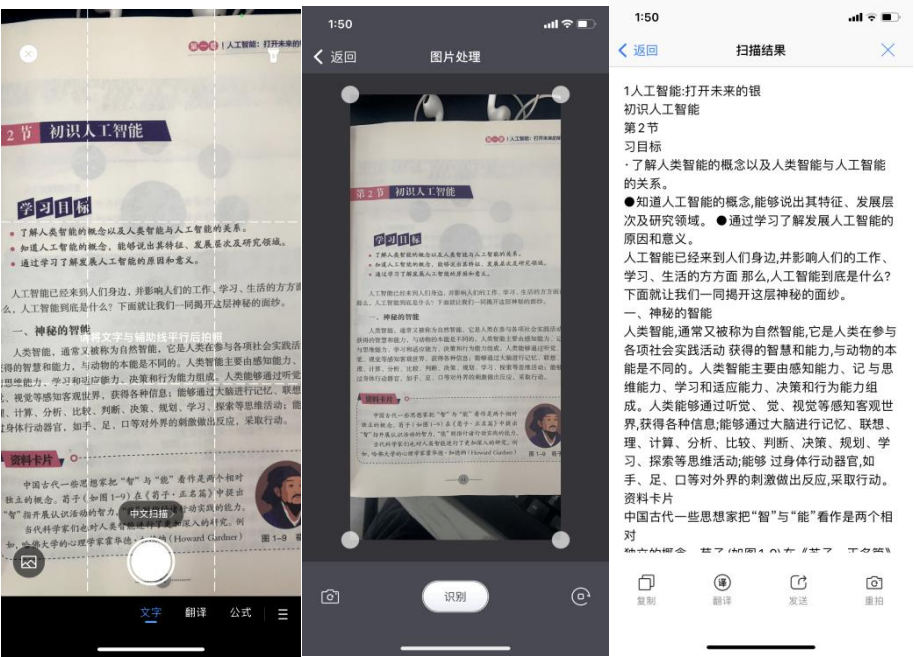


图 2. 1. 2 讯飞输入法拍照输入界面

表 2. 1. 2 计算机视觉体验任务记录表

项目	记录
体验的工具名称	
体验的功能	
错误数（如文字识别，所识别页面中错误的个数）	
错误率（错误数/总数）	

二、计算机视觉的概念

经过对计算机视觉的初体验，是不是突然发现日常生活中有许多应用原来就是用的人工智能的计算机视觉技术啊，那么计算机视觉到底是什么，计算机视觉包括哪些分类呢？

1.计算机视觉的概念

计算机视觉是计算机对图像进行采集、处理、分析和理解，以识别各种不同模式的目标和对象的技术。也就是让机器像人一样能“看”到这个世界，同时又能“读懂”这个世界。计算机视觉包括**图像采集、图像处理、图像识别和图像理解**四个阶段，图像识别侧重于识别图像中的模式，如是否是一条狗，是否是一朵花。图像理解则侧重于理解图像中的对象，如是什么狗，是什么花，图像中的对象呈现什么状态或画面描述了什么意义，类似于小学生所做的看图说话。图像处理是图像识别和图像理解的衍生品，也是图像识别和图像理解的重要基础，图像处理是为了更好的图像识别和图像理解。图像采集是实现了类似数码相机和摄像头的功能，实现图像的获取。目前计算机视觉的研究重点在图像识别和图像理解。

计算机视觉是感知智能的重要应用，是通过视觉的模式来感知图像，并实现感知世界，这也是人工智能从感知智能走向认知智能的基础，是获取信息的通道，同时也是通用人工智能发展过程中必须突破的基础技术。

目前，生活中对计算机视觉已经有了较多的应用，如手机人脸解锁、手机或电脑的指纹解锁、抓拍机动车违法、自助停车缴费系统、拍照批改作业等，这些应用已经非常成熟，已经可以实现日常应用。

2.计算机视觉的发展水平

为了达到日常应用的水平，学者们做了大量的探索和实践。一些学者为了激励该技术的进步，也举办了一些挑战赛，比较著名的是 **imagenet** 挑战赛，即 **ILSVRC**，该挑战赛由李飞飞发起，从 **2010** 年到 **2017** 年每年一届，主题定在物体定位、物体检测、视频物体检测三个方向。由于 **2017** 年参赛队伍表现优异，举办者宣布

该赛事完成了使命，相关任务已经被宣布完成挑战。在 2017 年的挑战赛中，参赛者的最低定位错误率达到 0.062263，分类错误率达到了 0.03413¹。也就是计算机系统已经能以近乎 94%的成功率对物体在图像中的位置进行定位，能够以 97%的成功率对物体进行分类，这已经明显达到或超越了人类的最高水平。

3.计算机视觉的核心问题

如同 ILSVRC 的挑战主题，计算机视觉目前研究的核心问题包括：图像分类、物体识别、语义分割、视频分析。图像分类是对识别图像中的物体，并归入不同的分类中，完成图像打标签，如识别出动物、植物、人等不同类别。物体识别是检测出物体在图像中的位置，识别出物体是什么，并实现打标签，如猫、狗、玫瑰花等。语义分割属于图像理解，是找到物体之间的关系和联系，如蝴蝶在花上面飞。视频分析则是让计算机视觉能够看懂视频中的内容，同样属于图像理解，如一只猫追着蝴蝶。

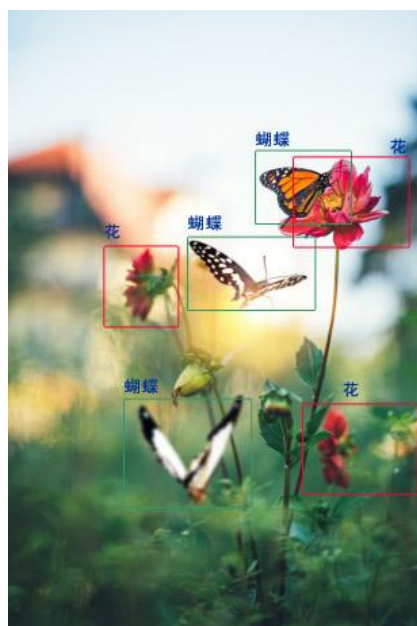


图 2.2.1 花与蝴蝶图像识别（图像来源：网络图库）

计算机视觉研究者的日常工作有点像在教计算机程序看图说话，在日常研究过程中，他们一直在探索如下问题的解决：

如在图 2.2.2 中：

(1) 这是不是一条狗？

¹ Imagenet.ILSVRC2017 results[DB/OL].<http://image-net.org/challenges/LSVRC/2017/results,20210421>.

- (2) 这是什么物体？
- (3) 狗在图中什么位置？
- (4) 图里都有什么？



图 2.2.2 宠物图像（图像来源：网络图库）

4.计算机视觉的一般过程

计算机视觉包括图像采集、图像处理、图像识别、图像理解等过程，在一般情况下，并不是所有的计算机视觉相关的应用覆盖全部过程，如有的应用专注于图像处理，有的应用专注于图像识别等，但一般认为图像识别是比较通用的过程和模式，下面就以图像识别为例介绍图像识别的一般过程。

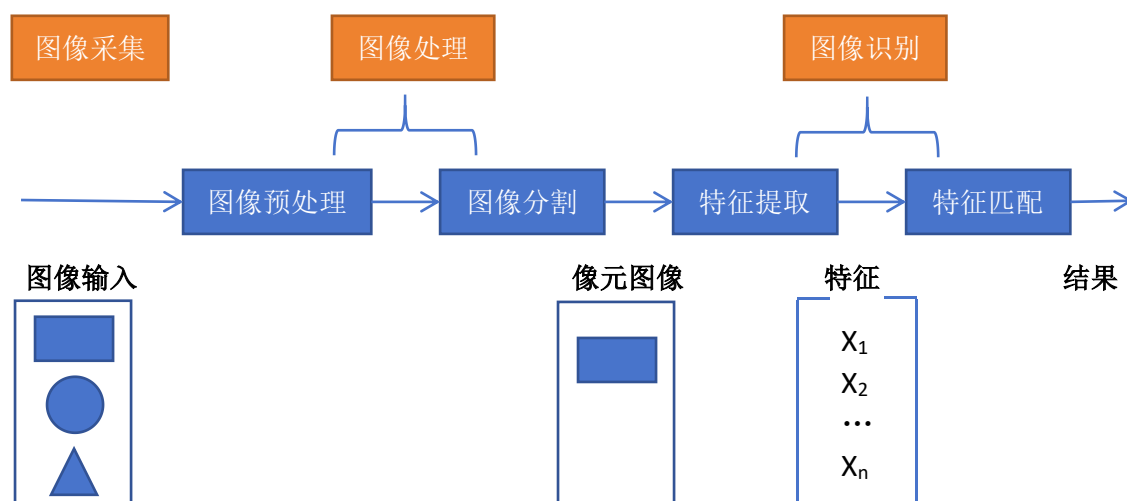


图 2.2.3 图像识别的一般过程

图像识别的一般过程包括，图像输入，图像预处理，图像分割，图像特征提取和图像分类。图像处理是为了消除图像像素、大小等因素的影响，实现图像的

同质化，便于进行分析和比较。图像特征提取是在归一化处理后的图像中，进行图像特征的发掘和提取，然后将特征与训练好的图像分类模型进行对比，即特征匹配，并计算图像属于相关分类的概率，并输出相应的概率值，从而实现图像识别的整个过程。

三、计算机视觉的应用

人工智能由于不会疲倦，而且部分能力已经开始超过了人类的最高水平，因此在计算机视觉方面，已经有大量的应用场景。如光学文字识别、物体识别、人脸识别等。

1. 光学文字识别

光学文字识别可以实现纸质文字的数字化，实现拍照输入功能，还可以实现拍照翻译，甚至可以实现实时场景翻译，数学公式识别、表格识别、pdf 或图片转 word、拍照批改作业等。还可以基于这些基础功能进行拓展应用，如拍照查生字等。光学文字识别应用案例见应用案例 1 和应用案例 2。

应用案例 2-2：利用讯飞输入法拍照识别公式

场景描述：数学公式的输入一直是一个“专业体力活”，无论是长期从事数学相关工作的专业人员，还是普通的学生，利用文字编辑工具输入公式都是比较头疼的，是否存在这样的工具，能够识别纸上的公式，从而提高公式输入的效率？在这样的情境下，拍照输入公式的需求便应运而生了。

体验步骤：

（1）在移动终端上，如苹果手机（IOS）、ipad、Android 手机、平板电脑，访问相应的 APP 市场，并下载讯飞输入法，以下步骤以 IOS 下的操作为例，其他系统略有不同；

（2）在讯飞输入法，“我的”，“工具箱”中启用拍照输入功能；

（3）在讯飞输入法，“我的”，“工具箱”中“拍照输入”点击“开始使用”，选到“公式”，按照提示开始使用；


（4）在使用后可以对比检测的正确率，并思考如下问题：

- 为什么需要拍照输入公式？它解决了什么问题？

- 该解决方案存在什么问题？你是否有其他更好的解决方案？

应用案例 2-3：拍照检查作业（小猿口算）

表 2.3.1 拍照检查作业应用信息

环境	工具名称	功能简介	获取方式
移动终端下 (IOS 和 Android)	小猿口算	拍照实现一秒检查小学作业。	1. 相应的市场输入小猿口算下载 2. 扫描以下二维码：  3. 访问网址： http://www.xiaoyuankousuan.com/

场景描述：你在家里的日常作业过程中，是否碰到一些不会的题目或者课堂中没听懂的题目，想要得到帮助，而你的家长没有办法进行辅导？抑或，你做了作业后，想及时的知道自己做的是否正确，并及时的订正错误，防止误区沉淀。而作业完成后一般需要第二天老师批改后才知道正确与否，此时，你早已忘记自己当时是怎么想的了，这无疑不利于纠正错误的方式和方法，是否有方法解决这样的问题呢？这时候就需要一个智能的检查作业小助手了，下面就以小猿口算为例体验一下拍照检查作业的强大功能。

体验步骤：

(1) 在移动终端上，如苹果手机 (IOS)、ipad、Android 手机、平板电脑，访问相应的 APP 市场，并下载小猿口算（为了防止作为初中生的你偷懒，我们选用小学版的此类软件来体验！），以下步骤以 IOS 下的操作为例，其他系统略有不同；

(2) 打开校园口算，点击拍照检查，对准你的作业，让手机与你的作业保持平行并拍照，完成 1 秒检查作业；

(3) 错题会自动进入自己的错题集，在考试之前可以打印错题进行复习。



图 2.3.1 拍照检查作业界面

(4) 完成体验后，请判断正确率如何？并思考如下问题：

- 拍照检查作业解决了什么问题？
- 小猿口算的模式是否存在缺陷或者不利影响？如何避免？
- 你是否有其他的更好的解决方案？

【拓展阅读】

我们已经完成了光学文字识别的体验，你是否觉得光学文字识别有用呢？它这么有用，而且识别准确这么高，他背后的原理和处理过程是什么呢？

(1) 图像输入：采集所要识别的图像，如名片、身份证、护照、行驶证、驾驶证、书本、作业等，一般是采用和数码相机同样的功能的摄像头进行图像采

集。

(2) 图像预处理：此过程包含二值化（将图像上的像素点灰度值设置为 0 或 255，也就是将整个图像呈现出明显的只有黑和白的视觉效果）、去噪、倾斜度矫正等。

(3) 版面分析：将所要识别的文档分段、分行处理。

(4) 字符切割：此时需要字符定位和字符切割，定位出字符串的边界，然后分别对字符串进行单个切割，单个分割出来的字符再做识别。

(5) 字符特征提取：提取字符特征，为后面的识别提供依据。

(6) 字符识别：将当前字符提取的特征向量与特征模板库进行模板粗分类和模板细匹配，识别出字符。

(7) 校正并输出：对文字按照上下文及语境进行校正，并按照文字原有排版输出为纯文本、word 或者 pdf 文件。²

由以上过程可知，光学文字识别的核心是文字特征提取和字符识别，一个高效率的特征提取算法和一个有效的字符特征模型是光学文字识别成功的关键。

2. 物体识别

物体识别，是识别图像中这是什么物体，属于什么种类，描述物体的位置、名称、分类等信息，如识别植物、动物、商品、建筑物、风景、动漫、食物等物体，还可以识别相关的场景，并返回场景中的物体。物体识别的核心是构建所需识别物体的模型，通过模型匹配实现模式识别。

如表 2.3.1 所示，提供了两种环境下的物体识别应用，可根据自己的信息化环境挑选应用进行探索，下面就以形色为例来体验物体识别中的花草识别。

表 2.3.2 物体识别应用

环境	工具名称	功能简介	获取方式
Windows 操作系统	百度 AI 开放平台 通用物体和场景识别	识别常见的物体和场景	1. 网址： https://ai.baidu.com/tech/imagerecognition/general 2. 二维码：

² 中央电化教育馆.人工智能（高中版）[M].北京：北京师范大学出版社,2019-09,136-138.

			
环境	工具名称	功能简介	获取方式
移动终端下 (IOS 和 Android)	形色	1 秒识别各种植物。	1. 相应的市场输入形色下载。 2. 或访问网址: http://www.xingseapp.com/ 3. 二维码: 

应用案例 2-4: 利用形色拍照识别植物

场景描述: 你是否对大自然充满了好奇? 在访学路上, 在小区公园, 在植物园里, 各式各样的植物, 琳琅满目, 你是否能准确的识别这些植物? 除非你自己就是非常厉害的植物爱好者, 或者有个非常聪明的植物学家在你旁边, 你可以随时向他提问, 不然我们只能“看看而已”, 是否有这样的工具, 能够实时的告诉我这是什么植物, 这种植物属于什么类别, 有什么用途呢? 这里我们来看一下“形色”这款 APP 是否能解决这个问题。

体验步骤:

(1) 在移动终端上, 如苹果手机 (IOS)、ipad、Android 手机、平板电脑, 访问相应的 APP 市场, 并下载形色, 以下步骤以 IOS 下的操作为例, 其他系统略有不同。

(2) 打开形色, 点击底部的相机图标。

(3) 将镜头对准要识别的对象, 点击拍照按钮, 1 秒识别拍照对象, 点击识别详情, 查看识别结果。如图 2.3.2 所示。

(4) 完成体验后, 请判断该软件的识别准确度如何? 并思考如下问题:

- 该应用解决了什么问题？你觉得该软件解决该问题的效果如何？
- 你觉得该软件在解决该问题时是否有缺陷？如何避免这些缺陷？
- 你是否还有更好的解决方案？

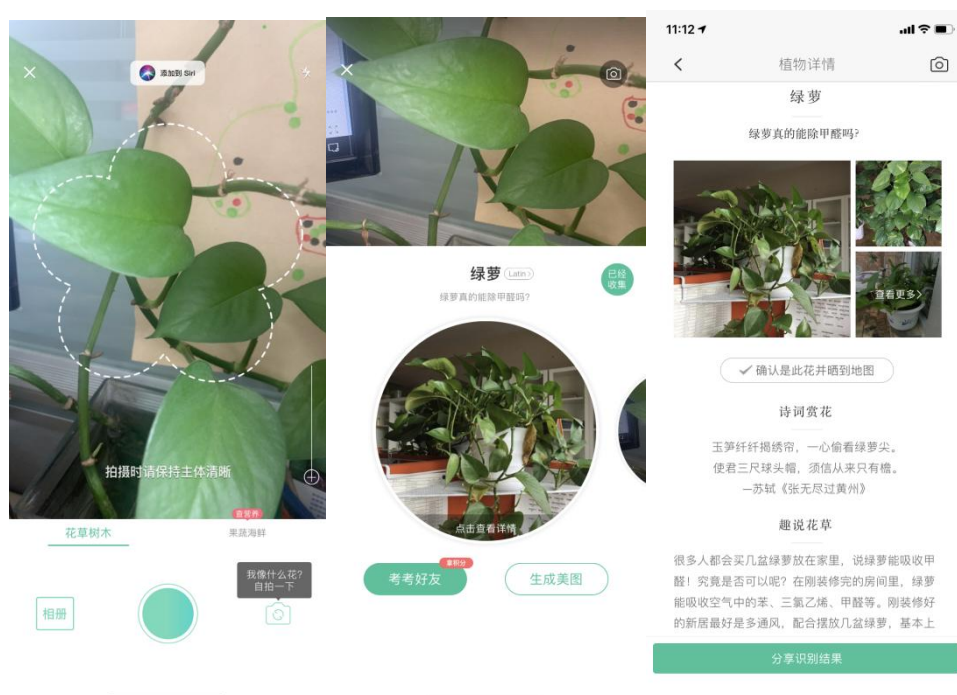


图 2.3.2 形色识别界面（IOS 版）

【拓展阅读】



图 2.3.3 百度 AI 开放平台的通用物体和场景识别体验³

经过物体识别的应用体验，你是否感受到了它的强大功能？物体识别的识别过程是什么？有哪些常见的物体识别方法呢？下面我们就一起来了解下吧！

● 物体识别的步骤

物体识别的步骤和图像识别的过程大同小异，首先是图像采集，实现图像的输入。然后是图片的预处理，图像处理是在不改变图像所存储的信息的前提下，实现不同图片之间的特征采集和匹配，是为了让每张图片的外观特征一致，如颜色模式、整体亮度、尺寸等。通常使用的图像预处理手段有：编码，阈值或者滤波运算，模式改善正规化，离散模式运算等。

接下来是特征提取，常见的图像特征如灰度、亮度、纹理和形状特征，还有描述图像的一些常见属性特征如颜色直方图、空间频谱等。

特征建模，完成特征提取后，需要对特征进行筛选和建模。在成熟的应用中，一般不需要进行建模工作，因为只需要将提取的特征与训练好的模型进行匹配即可。在新的物体识别任务中，需要首先建立所要识别的这类物体的模型，以便于下一步的特征模型匹配。之所以需要进行特征选择是因为图像本身的特征种类非常多，在不同的应用场景中，需要不同的特征集，还有就是如果采用原图像的特征全集，对系统的开销要求就非常高，系统计算的效率就会降低，而且不同类别

³ 百度 AI 开放平台.通用物体识别和场景识别[DB/OL].<https://ai.baidu.com/tech/imagerecognition/general>, 2021 年 4 月 22 日.

的物体识别并不需要全部特征就可以完成识别，因此一般会在不同识别场景中选择不同的图像特征集。选择好了图像特征集后，就是对图像特征集合进行建模，对特征与特征之间的空间结构关系进行描述，这样就形成了这类物体的通用模式——模型。

特征匹配。将所提取的物体特征与模型进行比对，计算匹配的概率。

最后一步，将匹配结果输出。⁴

● 物体识别的常见方法

物体识别的主要方法有很多，包括基于统计的方法、基于物体部件的方法、基于模型的物体识别等。目前比较流行的就是基于模型的物体识别方法，前述的物体识别的过程就是基于模型的物体识别的过程。这种方法的核心在于选取合适的图像特征并选择合适的特征匹配算法。

3. 人脸识别

人身上的特征信息有很多，在很多方面也已经开始应用，如指纹信息、人脸信息、声纹信息、虹膜信息等。以上应用，除了声纹信息，其余的都是图像识别的典型应用。我们以人脸识别为例进行阐述。人脸识别是一种基于人的脸部特征信息进行身份识别的生物特征识别技术，也称为人像识别、面部识别。由于人脸信息是人的独特标识，由于这种独特性，人脸信息属于个人隐私信息，在应用和开发人脸相关的程序时，要充分尊重个人隐私信息的保护。

为了充分保护人脸隐私信息，在本部分进行应用的人脸图像都使用工具生成，也就是不采用真人人脸做应用体验（可以使用百度 AI 开放平台的人脸搜索功能来检验）。生成不存在的人脸工具地址参加下表第一行。

⁴ 知乎.物体识别技术一览[EB/OL].<https://zhuanlan.zhihu.com/p/36668416>, 2021 年 4 月 22 日.

表 2.3.3 人脸识别应用

环境	工具名称	功能简介	获取方式
Windows 操作系统（有浏览器的终端都可以访问）	这个人不存在	刷新页面生成一个不存在的人的人脸图片。可以用于人脸识别训练和测试场景。	1. 网址： https://thispersondoesnotexist.com/ 2. 二维码： 
windows 操作系统（有浏览器的终端都可以访问）	百度 AI 开放平台 人脸检测与属性分析试用页面	人脸检测与属性分析。	1. 访问网址： https://ai.baidu.com/tech/face/detect 2. 二维码： 

应用案例 2-5：基于百度 AI 开放平台的人脸检测与属性分析

场景描述：人脸识别的应用目前主要在解锁、人脸考勤、安防等领域。那你的脸到底隐藏了哪些秘密呢？这些秘密能带给我们哪些信息？通过人脸能预测人的旦夕祸福是否靠谱？这些问题是我们在人脸识别领域的困惑，那么是否有应用来让我们破解这些问题呢？下面就以人脸检测与属性分析的功能试用网页来体验一下。

体验步骤：

- （1）浏览器打开这个人不存在网站，将页面中的人脸图像下载到本地。
- （2）浏览器打开上表第二行的百度 AI 开放平台人脸检测与属性分析页面，在功能演示部分，点击本地上传，将刚才下载的人脸图像选中并上传。
- （3）完成人脸识别、人脸关键点标注和相关属性值的计算与标注，如图

2.3.4 所示，可以标注年龄、性别、颜值、表情、眼镜、情绪、戴口罩、头部姿态等信息。

（4）完成体验后，请思考并回答以下问题：

- 该应用解决了什么问题？
- 这个问题是否很好的完成了解决？
- 你是否还有更好的解决方案？
- 你觉得这个应用还可以用在什么地方或场景中？

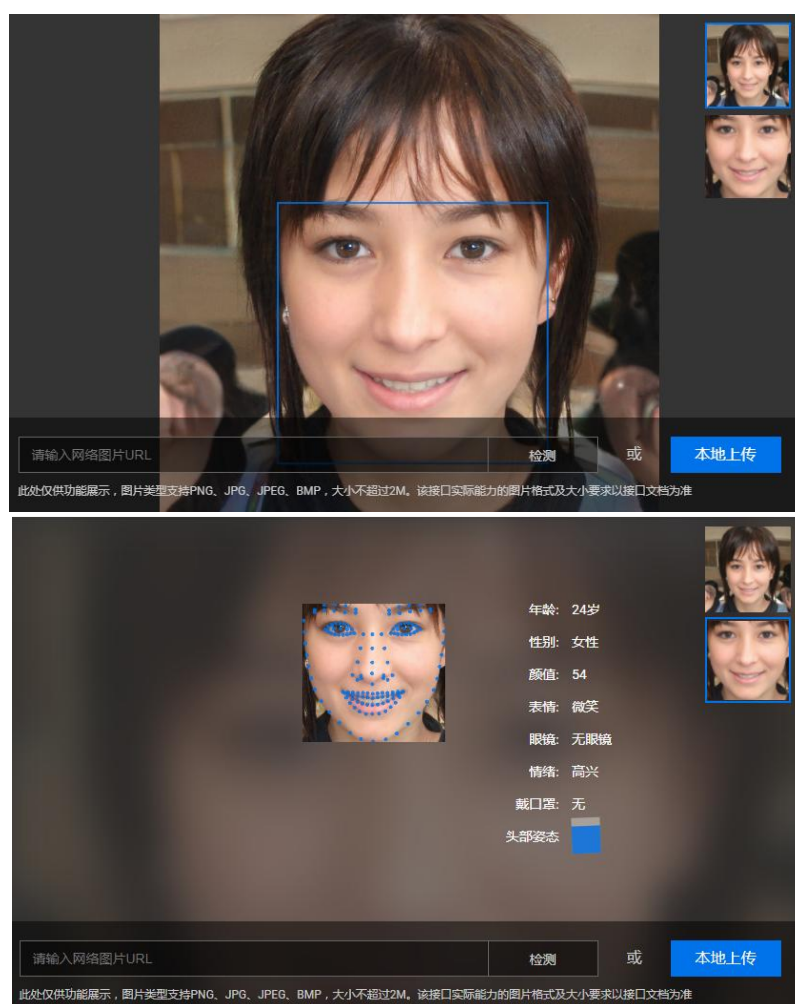


图 2.3.4 人脸检测与属性分析

【拓展阅读】

人脸识别是基于人的脸部特征信息进行身份识别的一种生物识别技术。目前，人脸识别技术已经非常成熟，在 3D 人脸识别方面也是有了发展，人脸识别的本质和物体识别相同，只是识别的对象是人脸这种特殊对象。

人脸识别的关键步骤包括：图像采集，图像预处理，人脸检测，人脸对齐，人脸编码，人脸匹配。其中，人脸检测、人脸对齐、人脸编码和人脸匹配是关键步骤。人脸检测是寻找图像中人脸的位置，并标注出人脸的位置信息，或者将人脸切割出来。一般检测和识别人脸的常用算法是方向梯度直方图（HOG），先将图片灰度化，计算图像中的像素的梯度，转成 HOG 格式，这样就获得了人脸的位置信息。人脸对齐是将不同角度的人脸图像对齐成同一种标准的形状。如将所有的眼睛、鼻子、嘴巴放在图像中的同一个位置，以便于分析这些位置的图像特征。人脸编码是将人脸图像的像素信息转换为特征向量，不同的算法转换模式不近相同，但本质都相同，在同一个算法下，同一个人的人脸特征的向量应该是相同的。人脸匹配是将特征向量与模型进行比对，得出人脸与该模型匹配的概率。

人脸识别算法的发展过程，经历了传统机器学习和深度学习两个大的阶段，在传统机器学习中，有几何特征法、表象特征法和纹理特征法。在深度学习中，普遍使用卷积神经网络建立模型，从输入图像中提取特征，通过计算余弦距离法来进行人脸识别。在深度学习算法中，学者们不断的在人工神经网络的宽度、深度、数据集的规模等方面进行探索，以提升算法的执行效率，实现网络的轻量化。

人脸识别目前有了广泛应用，主要使用人脸识别来鉴别身份，然后按照身份提供不同的服务和判断不同的职能权限。如人脸门禁、人脸支付、无人超市、无人业务办理（如自主开设银行账户）、娱乐等。

四、计算机视觉的开发




经过对计算机视觉技术的感知、应用等体验，你是否对计算机视觉有了初步的理解？充分应用计算机视觉技术，可以让我们实现原本要靠人的视觉才能完成的任务。掌握了相关知识后，你是否想开发一个属于自己的程序呢？你肯定跃跃欲试了吧！

下面，我们按照图形化编程和代码编程两种模式来介绍一下计算机视觉下面

的编程案例，希望你通过这些案例的学习，能够掌握相应环境下开发软件的一般步骤，并能够开发出你自己的计算机视觉应用程序。

1.图形化编程模式

表 2.4.1 图形化编程工具档案

环境	工具名称	功能简介	获取方式
跨终端支持，支持 windows、Mac、IOS、Android 等系统	慧编程	一款国内厂家自主研发的图形化编程软件，还支持 Python 编程，支持机器人编程和人工智能编程，还可以连接相关的硬件进行硬件编程。是一款比较综合的适合进行人工智能编程的工具软件。	1. Windows 和 Mac 版下载网址： https://mblock.makeblock.com/zh-cn/download/ 2. 移动终端扫描二维码：  Android 扫码下载 支持安卓 6.0 及以上 (仅限 Arm 架构，不支持 X86 平台)  iOS 扫码下载 支持 iOS 10.0 及以上
支持 windows、Mac、Linux 等系统	Mind+	Mind+是一款拥有自主知识产权的国产青少年编程软件，集成各种主流主控板及上百种开源硬件，支持人工智能（AI）与物联网（IoT）功能，既可以拖动图形化积木编程，还可以使用 Python/C/C++等高级编程语言，让大家轻松体验创造的乐趣。	1. 下载网址： http://mindplus.cc/download.html 2. 扫码查看下载页： 

前驱知识准备：图形化编程的界面布局；知道角色背景造型等基本概念；图形化编程的事件（当绿旗被点击）、外观（说**2 秒）、控制（等待 1 秒、重复执行、如果那么）等基本积木；知道图形化编程的基本方法。

硬件准备：带有摄像头和麦克风的笔记本电脑（含 Mac 笔记本）、台式机电脑。注：平板电脑暂不支持机器学习模块。慧编程 mBlock5.3.0 支持 Windows7、Windows10、macOS 10.12+版本的操作系统，推荐 64 位操作系统。

软件工具准备：参照表 2.4.1 所示第一行的获取方式下载并安装慧编程工具。
完成账号注册并登陆（可以应用更多功能）。

注册方式：

- （1）打开 mBlock v5.3.0。
- （2）点击右上角头像，如图 2.4.1 所示。
- （3）点击注册，输入邮箱和密码。
- （4）点击获取验证码。
- （5）登陆邮箱收取验证码，返回 mBlock 输入验证码。
- （6）点击注册完成注册。返回登陆页面登陆即可。

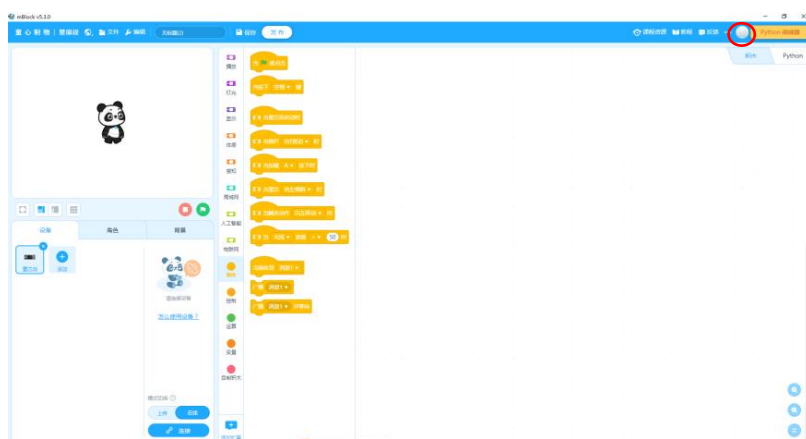


图 2.4.1 慧编程 mBlock v5.3.0 主界面

图 2.4.2 慧编程注册界面

任务：搭建一个能识别人的年龄、性别和是否高兴的魔镜

任务参考步骤：

- (1) 打开慧编程 mBlock 程序，并点击右上角头像登录。如图 2.4.1。
- (2) 点击舞台下方的“角色”，点击积木类别区下方的“添加扩展”，如图 2.4.3，将扩展区的第一个“认知服务”添加到积木区，如图 2.4.4 和图 2.4.5。

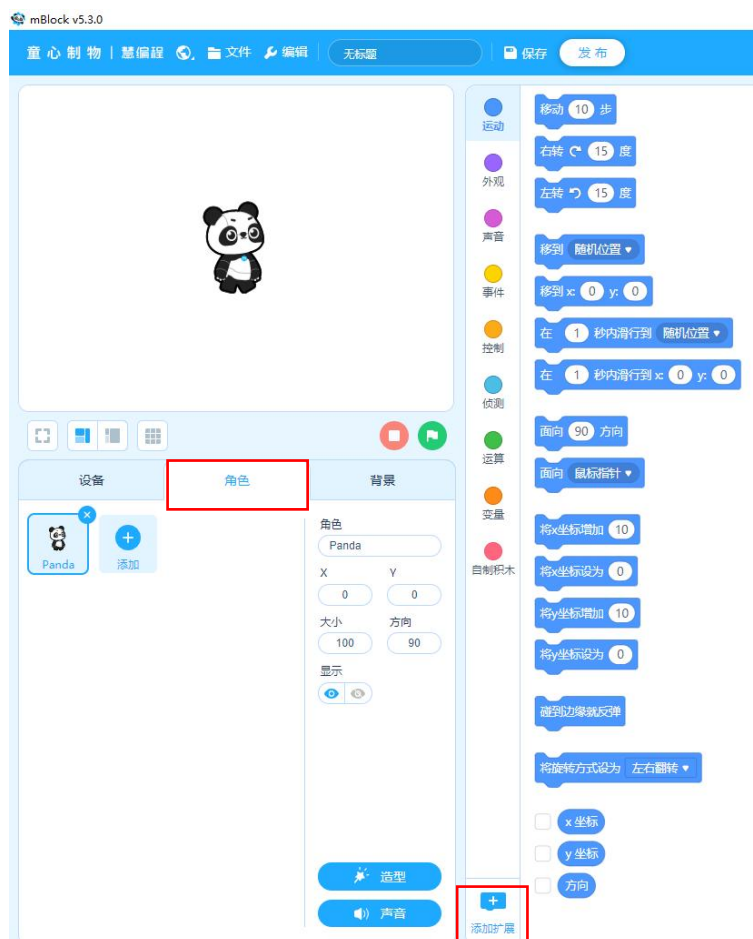


图 2.4.3 添加扩展界面

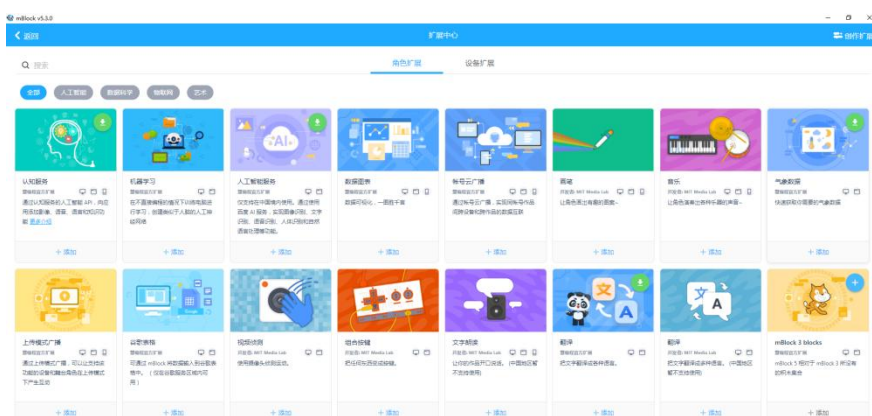


图 2.4.4 角色扩展界面

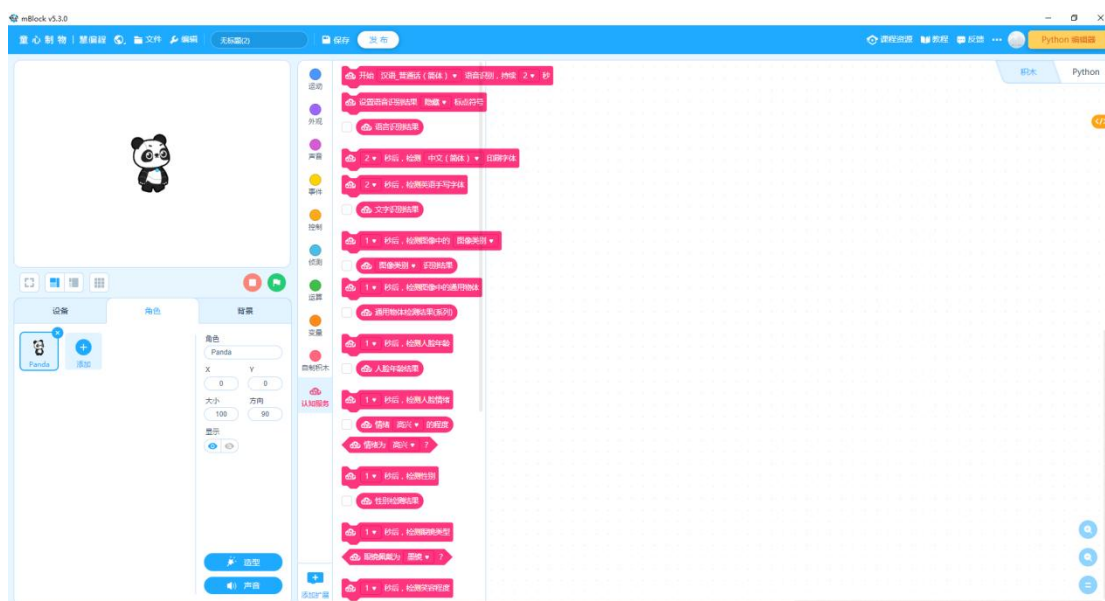


图 2.4.5 认知服务积木块

(3) 识别性别，并让熊猫（Panda）角色说出你的性别。



图 2.4.6 识别性别并说出

(4) 识别性别提高任务：识别性别结果，并用中文说出，如果为男的说“你好帅哥！”，如果为女的说“你好美女”，说的持续时间为 4 秒。



图 2.4.7 带判断的性别识别

(5) 接着识别年龄，并说出年龄值。思考：和性别识别相比年龄识别结果

是否准确？你觉得影响识别结果的因素有哪些？积木程序如图 2.4.8 所示。



图 2.4.8 识别人的年龄并说出年龄识别结果

(6) 检测人的情绪，如果开心说“你好开心啊！”，如果生气说“不要生气嘛！”。



图 2.4.9 魔镜完整程序

拓展任务 1（难度：简单）：搭建一个光学文字识别的应用，能够识别你放在摄像头前的书的名字。

拓展任务 2（难度：中等）：口算练习检测系统。系统给出一道题目，你用纸填写答案，程序判断对错。如系统出题 1020×31 ，在 10s 内在纸上写出答案，并展示在摄像头前，程序判断对错，对了说“正确”，错了说“错误”。

拓展任务 3（难度：较难）：搭建一个水果识别系统，能够识别香蕉、苹果和小番茄。提示：用到机器学习拓展积木，用机器学习训练程序识别三种水果。

2. Python 编程模式

前驱知识准备：编程的基本模式如 IPO 模式，知道输入、处理、输出的概念；Python 的基本语法：运算符与表达式、控制流（顺序、分支、循环）、输入与输出，库安装和库引用，Python 的编程环境配置（如 Anaconda 或者单独的 Jupyter notebook）、安装与使用。

硬件准备：带有摄像头和麦克风的笔记本电脑（含 Mac 笔记本）、台式机电脑。

软件准备：Python 3.8.5 及以上版本。直接下载 Python 安装包或者通过 Anaconda 安装时一起安装 Python。

表 2.4.2 Python 编程环境档案

环境	工具名称	功能简介	获取方式
跨终端支持，支持 window s、Mac 等系统	Python	简单易学的编程语言。2021 年 4 月的最新版为 3.9.4（以下案例都基于该版本）。	Windows 和 Mac 版下载网址： https://www.python.org/downloads/
支持 window s、Mac、Linux 等系统	Anaconda	Anaconda 是 Python 的一个集成开发环境，程序包括 Conda、Python 以及一大堆安装好的工具包，比如：numpy、pandas 等，可以方便的安装扩展库及相关的依赖库。	网址： https://www.anaconda.com/products/individual

● 基于 cnocr 的中文光学文字识别（不用摄像头版本）

（1）在保证 Python 环境具备的情况下，安装 cnocr 工具包。安装方法：在命令行模式，输入：`pip install cnocr` 或者在 Anaconda 环境下打开 CMD.exe Prompt 或者 Powershell Prompt，输入：`conda install cnocr`

（2）编写代码。

将需要识别的带有文字的图片放在与 python 程序相同的路径下。

然后编写以下代码识别文字：

```
from cnocr import CnOcr
ocr = CnOcr()
res = ocr.ocr('helloworld.jpg')
print("Predicted Chars:", res)
```



```
===== RESTART: C:/Users/Administrator/Desktop/shibieocr.py =====
Predicted Chars: [['H', 'e', 'l', 'l', 'o', ' ', 'w', 'o', 'r', 'l', 'd', '!', ' ', '你', '好', '世', '界']]
```

图 2.4.10 输出结果

● 基于 openCV 的人脸识别

(1) 安装相应的库，本案例需要用的库包括 numpy、opencv，安装方式为：在命令行模式下输入 `pip install opencv-python`。或者在 Anaconda 环境下打开 CMD.exe Prompt 或者 Powershell Prompt，输入：`conda install oepncv`

(2) 编写代码。参考代码如下：

代码①：从图片中识别人脸并标识

```
# -*- coding: utf-8 -*-
"""
Created on Wed Apr 28 09:53:05 2021

@author: Administrator
"""
import cv2
filepath="face001.jpg"
img=cv2.imread(filepath)
#读取图像
gray=cv2.cvtColor(img,cv2.COLOR_BGR2GRAY)
#图像转灰色
#cv2.imshow("Image",gray)

#x=y=10#坐标
#w=100 #矩形大小宽和高
#color=(0,0,255)
#cv2.rectangle(img, (x,y), (x+w,y+w), color,1)#绘制矩形
#cv2.imshow("Image",img)
#cv2.waitKey(0)
#cv2.destroyAllWindows()

classifier=cv2.CascadeClassifier("haarcascade_frontalface_default.xml")
#调用内置的人脸识别训练模型 下载地址: https://github.com/opencv/opencv/tree/master/data/haarcascades
color=(0,255,0)#定义绘制颜色，R、G、B，颜色范围0-255

faceRects=classifier.detectMultiScale(gray,scaleFactor=1.3,minNeighbors=3,minSize=(32,32))
#调用识别人脸
print(faceRects)
if len(faceRects):
    #大于0则表示检测到了人脸
    for faceRect in faceRects:
        x,y,w,h=faceRect
        #将每一个人脸的坐标和宽度高度值分别赋值给x.y.w.h
        cv2.rectangle(img, (x,y), (x+w,y+h), color,2)
        #框出人脸
cv2.imshow("image",img)
#呈现图像
c=cv2.waitKey(10)
cv2.waitKey(0)
cv2.destroyAllWindows()
```

图 2.4.11 识别并标识人脸源代码

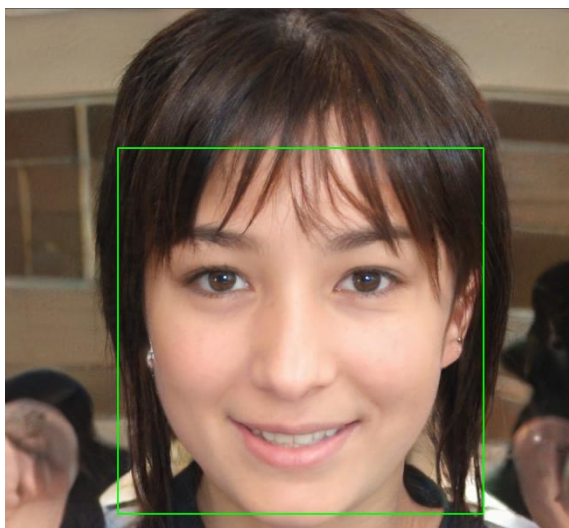


图 2.4.12 人脸识别检测效果（注：人脸图像来自 thispersondoesnotexist.com）

代码②：从摄像头中识别人脸并标识

```
# -*- coding: utf-8 -*-
"""
Created on Wed Apr 28 09:53:05 2021
@author: Administrator
"""

import cv2

def recogfacefrompic(img):
    gray=cv2.cvtColor(img,cv2.COLOR_BGR2GRAY)
    cap=cv2.CascadeClassifier("haarcascade_frontalface_default.xml")
    faceRects=cap.detectMultiScale(gray,scaleFactor=1.2,minNeighbors=3,minSize=(50,50))
    if len(faceRects):
        print("共找到{0}个人脸".format(len(faceRects)))
        for faceRect in faceRects:
            x,y,w,h=faceRect
            cv2.rectangle(img,(x,y),(x+h,y+w),(0,255,0),2)
    else:
        print("没有检测到人脸！")
    cv2.imshow("Image",img)
cap=cv2.VideoCapture(0)
while (1):
    ret,img=cap.read()
    #cv2.imshow("Image",img)
    recogfacefrompic(img)
    if cv2.waitKey(1)&0xFF==ord('q'):
        break
cap.release()
cv2.destroyAllWindows()
```

图 2.4.13 从摄像头中检测人脸源代码

拓展任务 1（难度：中等）：通过摄像头实时识别纸质文字并输出。

拓展任务 2（难度：中等）：在识别人脸的基础上识别人的眼睛和嘴巴并标识出来。

拓展任务 3（难度：难）：识别是否佩戴口罩，并提醒。

思考与练习

1. 什么是计算机视觉？计算机视觉的常见应用有哪些？
2. 应用计算机视觉会不会导致个人隐私数据的泄露？如果会，该如何避免？
3. 请使用合适的工具或编程语言，搭建一个水果检测识别系统，至少能够识别香蕉、苹果和桔子三种水果。