卷积神经网络与计算机视觉

顾立平

01卷积和互相关

- 卷积和互相关都是信号处理中常用的运算方式。

$$(k \star f)(t) = \int_{-\infty}^{\infty} f(s)k(-s+t)ds$$

$$= \int_{-\infty}^{\infty} f(-s+t)k(s)ds$$

✓卷积是指两个函数的乘积在空间上的积分，它可以用来描述两个函数之间的关系，例如滤波器和信号之间的关系。

✓互相关则是指两个函数之间的相似性，它是将一个函数沿时间轴移动并与另一个函数相乘后再积分得到的结果。

01卷积和互相关

$$(k \star f)(n) = \sum_{s=-\infty}^{\infty} f(s)k(-s+n)$$

$$= \sum_{s=-\infty}^{\infty} f(-s+n)k(s)$$

- 在实际应用中，卷积和互相关广泛用于图像处理、语音识别等领域。

$$(k * f)(t) = \int_{-\infty}^{\infty} f(s)k(s+t)ds$$

$$= \int_{-\infty}^{\infty} f(s-t)k(s)ds$$

02变换不变性和变换方差

- 变换不变性和变换方差是机器学习中常用的概念。

✓变换不变性指的是某个变换对模型的影响是否能够被抵消掉，也就是说，变换后的模型是否与原始模型具有相同的性质。

✓变换方差则指的是模型对于不同的变换具有不同的敏感程度，也就是说，变换后模型的变化程度与原始模型的变化程度相比有何不同。

$$trans_a(k) \star f = k \star trans_a(f) = trans_a(k \star f)(t)$$

- 这些概念通常用于评估模型的鲁棒性和泛化能力，在选择合适的特征提取方法时也非常重要。

03常空间中的卷积是频率空间中的乘积

- 常空间中的卷积是频率空间中的乘积是一个常见的结论。
- 在频域中，两个函数的卷积等于它们的频谱相乘，这被称为傅里叶变换的基本性质之一。这个结论可以应用于许多领域，包括数字信号处理、图像处理和物理学等领域。
- 在实际应用中，可以通过将信号从时域转换到频域来利用这个性质来实现高效的信号处理和滤波操作。

04从系统设计的角度看卷积

- 从系统设计的角度来看，卷积是一种广泛应用于信号处理的技术。
- 在设计和分析简单系统时，如果满足线性和时间不变性的约束条件，那么卷积运算是不可避免的。
- 这种技术可用于处理与时间和空间相关的信号，例如电信号、音频信号和图像等。
- 在实际应用中，卷积神经网络已成为计算机视觉、机器音频处理和其他人工智能应用的重要组成部分。

05线性和平移不变系统的卷积和脉冲响应

- 在线性和时间不变系统中，卷积运算是描述系统响应的一种方式。当输入信号经过系统后，会产生输出信号，而这个输出信号可以通过输入信号和系统响应的卷积来计算得出。

$$H(ax) = aH(x) = ay$$

- 其中，系统响应指的是系统对单位脉冲响应的特性，也就是系统在接收到一个单位脉冲信号后的响应。通过测量系统响应，可以推断出系统对任何输入信号的响应。这种方法被广泛应用于信号处理、图像处理等领域。

$$H(x_1 + x_2) = H(x_1) + H(x_2) = y_1 + y_2$$

$$H(\delta(k)) = h(k)$$

05线性和平移不变系统的卷积和脉冲响应

$$x(k)\delta(n-k)$$

$$x(n) = \Sigma_{k=-\infty}^{\infty} x(k)\delta(n-k)$$

$$y(n) = (x \star h)(n)$$

$$\begin{aligned} H(x(n)) &= H\left(\sum_{k=-\infty}^{\infty} x(k)\delta(n-k)\right) \\ &= \sum_{k=-\infty}^{\infty} x(k)H(\delta(n-k)) \\ &= \sum_{k=-\infty}^{\infty} x(k)h(n-k) \\ &= (x \star h)(n) \\ &= y(n) \end{aligned}$$

06卷积和一维离散信号

$$(k * x)(n) = \sum_{s=-\infty}^{\infty} x(s)k(s+n)$$

$$= \cdots + x(-1)k(-1+n) + x(0)k(n) + x(1)k(1+n) + x(2)k(2+n) + \cdots$$

- 卷积是一种数学运算，用于描述两个函数之间的关系。
- 在一维离散信号的情况下，卷积是指将一个信号和另一个称为卷积核的信号进行逐点乘法和累加操作的过程。
- 这样得到的结果就是两个信号的卷积。卷积运算在数字信号处理中有广泛应用，如图像处理、语音识别等。

06卷积和一维离散信号

$$
\begin{aligned}
(k \star x)(-4) &= x_4 k_0 \\
(k \star x)(-3) &= x_3 k_0 + x_4 k_1 \\
(k \star x)(-2) &= x_2 k_0 + x_3 k_1 + x_4 k_2 \\
(k \star x)(-1) &= x_1 k_0 + x_2 k_1 + x_3 k_2 \\
(k \star x)(0) &= x_0 k_0 + x_1 k_1 + x_2 k_2 \\
(k \star x)(1) &= x_0 k_1 + x_1 k_2 \\
(k \star x)(2) &= x_0 k_2
\end{aligned}
$$

07卷积和二维离散信号

$$(k*x)(m,n)=\Sigma_{q=-\infty}^{\infty}\Sigma_{s=-\infty}^{\infty}x(m+q,n+s)k(q,s)$$

- 卷积和二维离散信号是指在一个二维网格上的两个函数之间的卷积运算。这个过程涉及到对两个函数的所有元素进行逐个相乘和求和的操作，从而生成一个新的函数。
- 在数字图像处理中，卷积运算是非常常见的，可以用来实现图像平滑、边缘检测、特征提取等操作。

07卷积和二维离散信号

$$A \star K = \begin{pmatrix} a_{00} & a_{01} & a_{02} & a_{03} \\ a_{10} & a_{11} & a_{12} & a_{13} \\ a_{20} & \boxed{a_{21}} & a_{22} & a_{23} \\ a_{30} & a_{31} & a_{32} & a_{33} \end{pmatrix} \star \begin{pmatrix} k_{00} & k_{01} & k_{02} \\ k_{10} & \boxed{k_{11}} & k_{12} \\ k_{20} & k_{21} & k_{22} \end{pmatrix}$$

$$Z = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix}$$

07卷积和二维离散信号

$$A_{padded} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_{00} & a_{01} & a_{02} & a_{03} & 0 \\ 0 & a_{10} & a_{11} & a_{12} & a_{13} & 0 \\ 0 & a_{20} & a_{21} & a_{22} & a_{23} & 0 \\ 0 & a_{30} & a_{31} & a_{32} & a_{33} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- 过滤图像指的是使用不同的滤波器对图像进行处理，以达到特定的目的。比如，可以使用高斯滤波器来实现图像平滑，去除噪声；也可以使用Sobel算子来进行边缘检测，找到图像中的轮廓等等。
- 不同的滤波器有不同的特点和适用范围，选择合适的滤波器可以提高图像处理的效果。

08过滤图像

Operation	Kernel ω	Image result g(x,y)
Identity	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	
Edge detection	$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$	
	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	
	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	
Sharpen	$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}$	

Replace pixel, say, a_{22}, with a_{11}-a_{13}-a_{31}+a_{33}
If there are no edges, this process will turn a lot of pixels off (black).

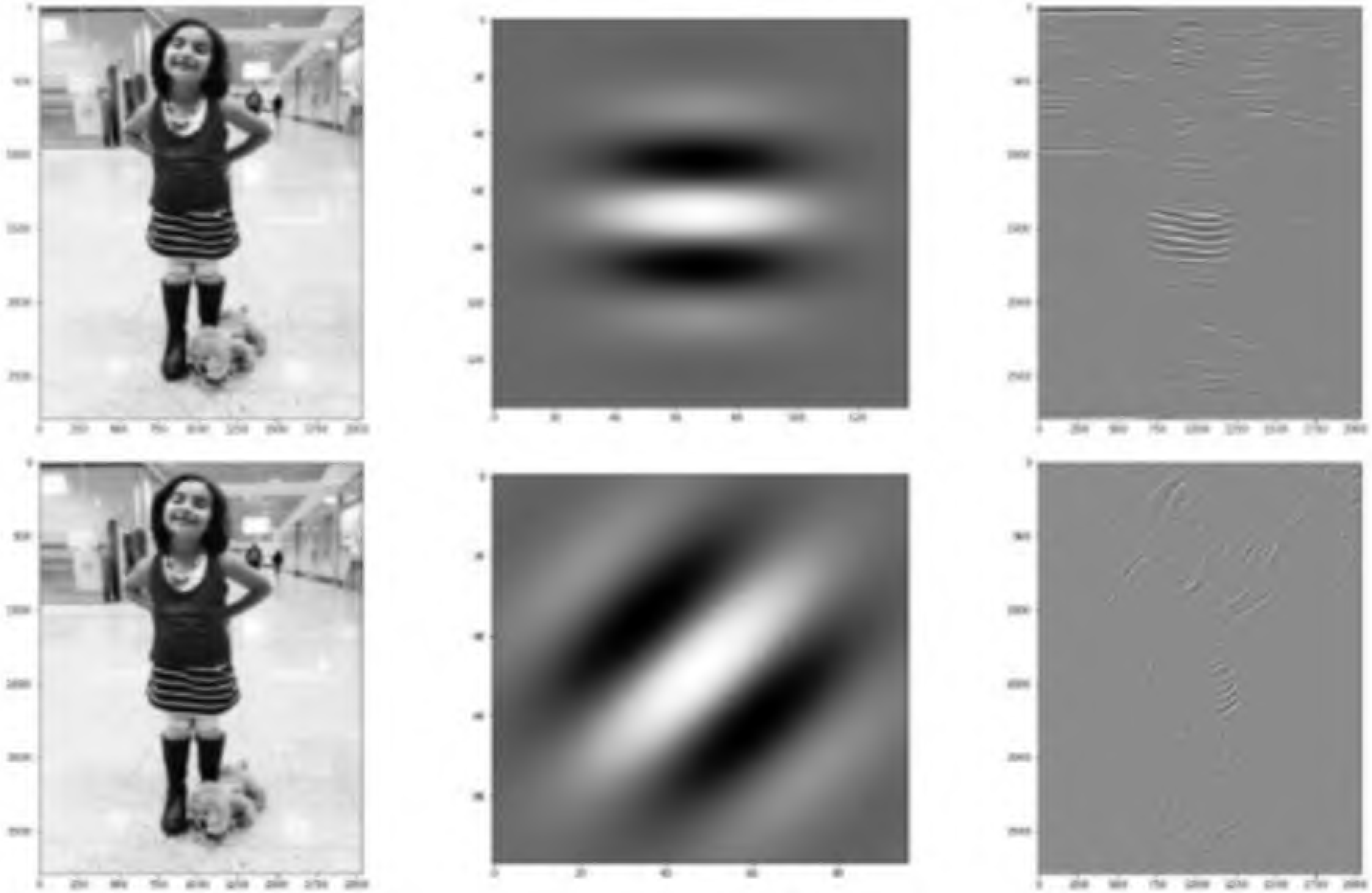
Make pixel very intense, and then subtract the sum of those vertically and horizontally around it.

08过滤图像

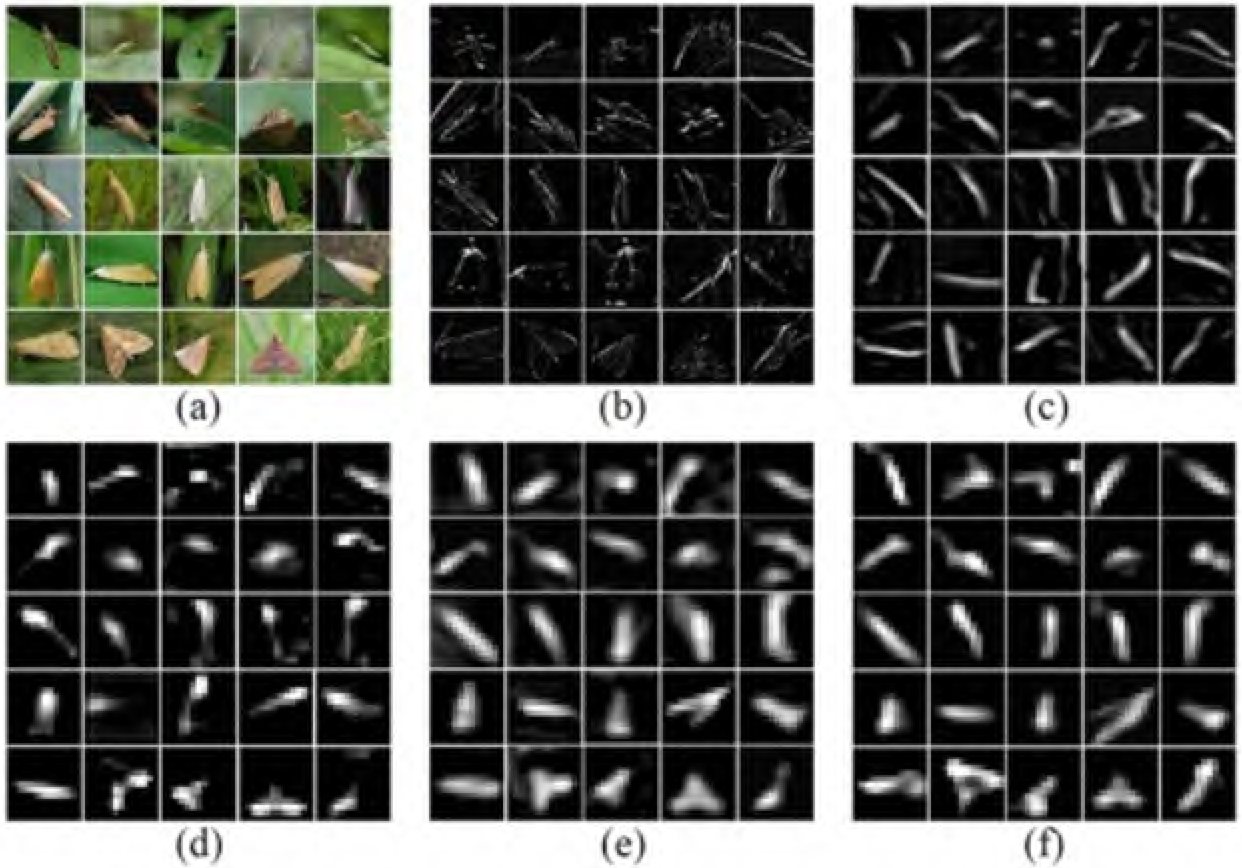
Box blur (normalized)	$\frac{1}{9}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$	
Gaussian blur 3 × 3 (approximation)	$\frac{1}{16}\begin{bmatrix}1&2&1\\2&4&2\\1&2&1\end{bmatrix}$	
Gaussian blur 5 × 5 (approximation)	$\frac{1}{256}\begin{bmatrix}1&4&6&4&1\\4&16&24&16&4\\6&24&36&24&6\\4&16&24&16&4\\1&4&6&4&1\end{bmatrix}$	
Unsharp masking 5 × 5 Based on Gaussian blur with amount as 1 and threshold as 0 (with no image mask)	$\frac{-1}{256}\begin{bmatrix}1&4&6&4&1\\4&16&24&16&4\\6&24&-476&24&6\\4&16&24&16&4\\1&4&6&4&1\end{bmatrix}$	

→ Each pixel becomes the average of all those around it, hence a blurry image.

08过滤图像



09功能图

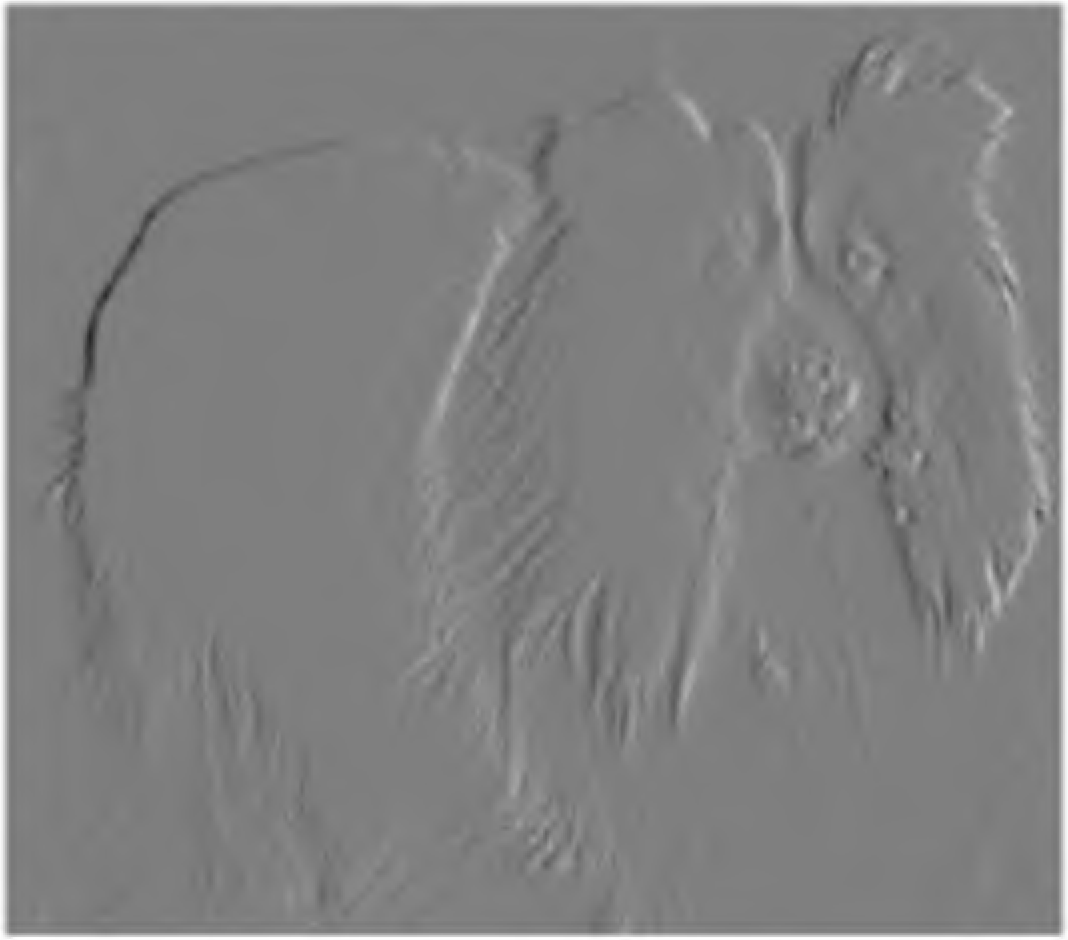


- 功能图是指一种可视化工具，用于展示系统、软件或者应用程序中的各个组件及其相互关系。通过功能图，我们可以更加直观地了解系统的架构和工作流程，方便进行设计、开发、测试等各个环节的工作。常见的功能图包括UML类图、状态图、时序图等等。

10线性代数符号

- 线性代数符号是一种数学符号和记号体系，用于描述和分析线性代数中的向量、矩阵和其他相关概念。
- 这些符号包括向量箭头、矩阵乘法符号、转置符号、加减号、标量乘法符号等等。
- 通过使用线性代数符号，我们可以更加清晰地表达和计算线性代数中的各种问题和方程式，从而更好地理解和解决实际问题。

10线性代数符号

11一维情形：Toeplitz矩阵的乘积

$$Toeplitz = \begin{pmatrix} k_0 & k_1 & k_2 & 0 & 0 & 0 & 0 \\ 0 & k_0 & k_1 & k_2 & 0 & 0 & 0 \\ 0 & 0 & k_0 & k_1 & k_2 & 0 & 0 \\ 0 & 0 & 0 & k_0 & k_1 & k_2 & 0 \\ 0 & 0 & 0 & 0 & k_0 & k_1 & k_2 \end{pmatrix}$$

$$(Toeplitz)x^t = k \star x$$

- 一维情况下的Toeplitz矩阵乘积是指将一个Toeplitz矩阵与一个一维信号相乘，得到的结果是一维卷积的精确结果。
- 在这个过程中，滤波器在信号上进行了滑动操作。

12二维情形：二重块的乘法

- 在二维情况下，二重块循环矩阵可以用来表示卷积运算。
- 该矩阵是由四个相同的元素组成的，其中两个位于主对角线上，另外两个位于副对角线上。这个矩阵可以用来执行二维卷积运算，其过程类似于在一维情况下使用Toeplitz矩阵进行卷积运算的过程。

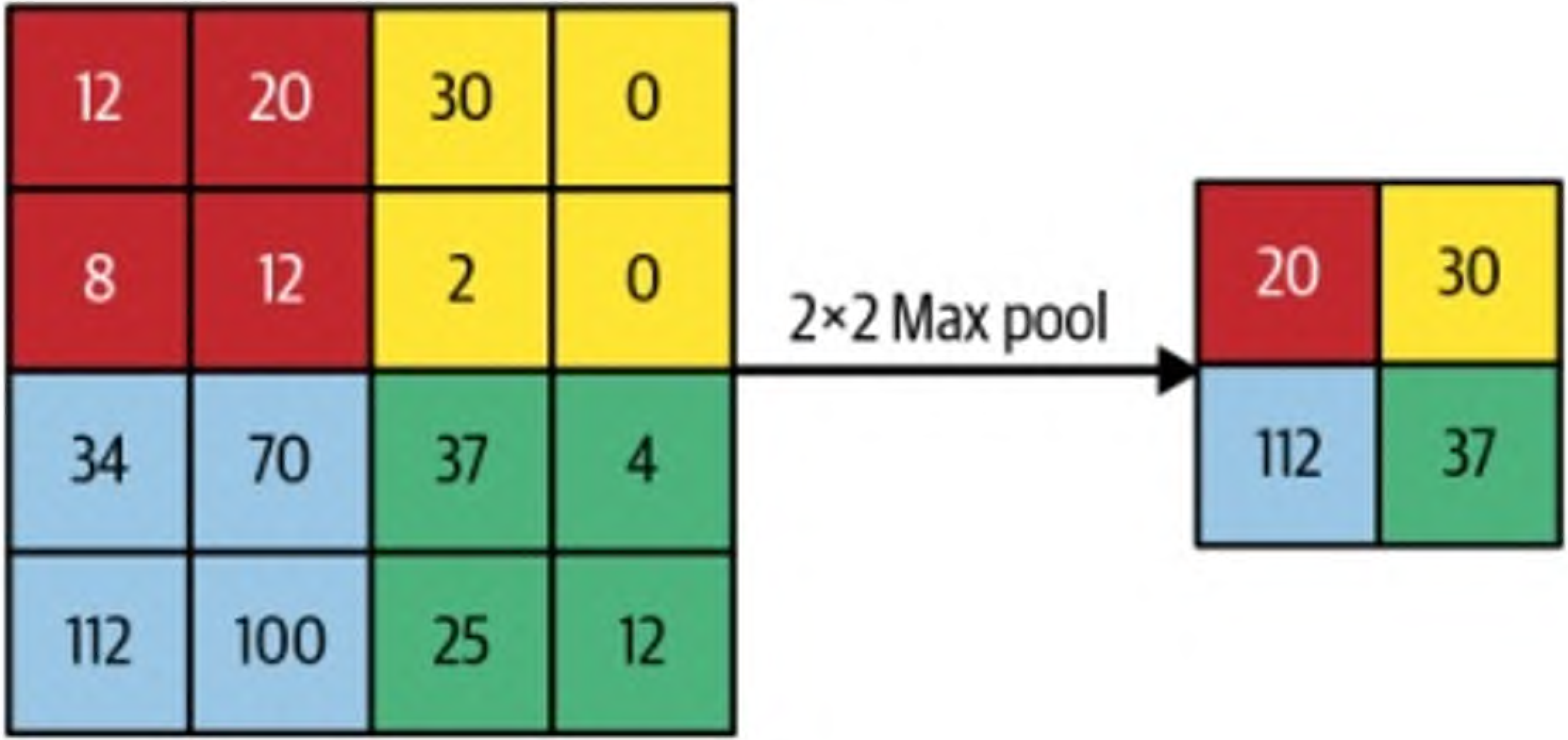
13循环矩阵

- 循环矩阵是指由一组具有循环性质的元素构成的矩阵。在本文中，循环矩阵被用于表示卷积运算，特别是在二维情况下，使用双块循环矩阵可以更有效地表示卷积运算。
- 循环矩阵的特殊结构使得它们能够更好地应用于卷积运算和其他相关操作中。

14池化

- 池化是一种常见的卷积神经网络层，用于减少特征映射的尺寸并提高模型的鲁棒性和计算效率。
- 池化操作通过对输入特征映射的一小部分进行聚合来生成单个输出值，例如取其中的最大值或平均值。
- 这样可以降低特征映射的空间分辨率，从而减少后续层的参数数量和计算量，同时还可以使特征更加鲁棒，对输入的微小变化不敏感。

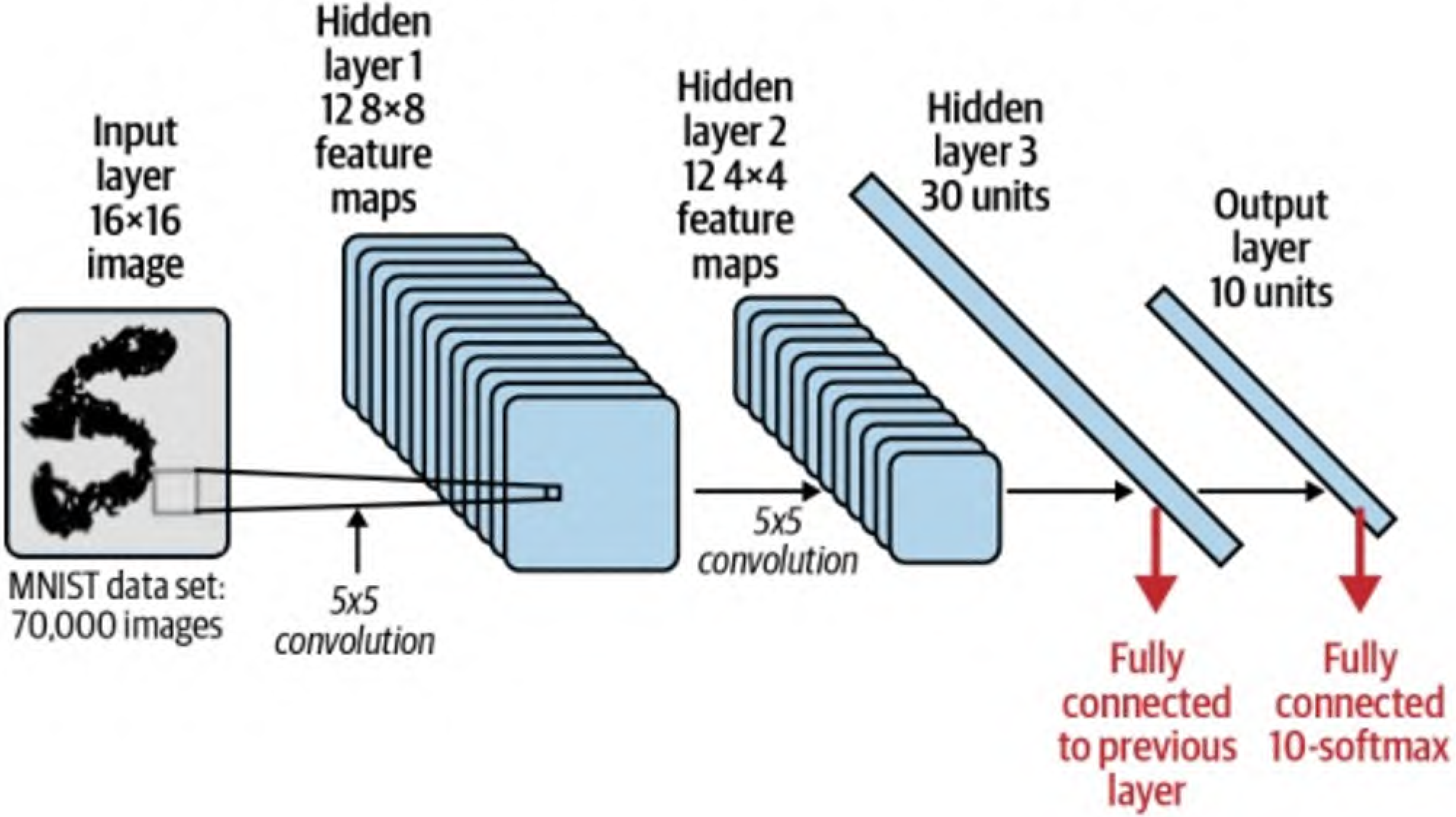
14池化



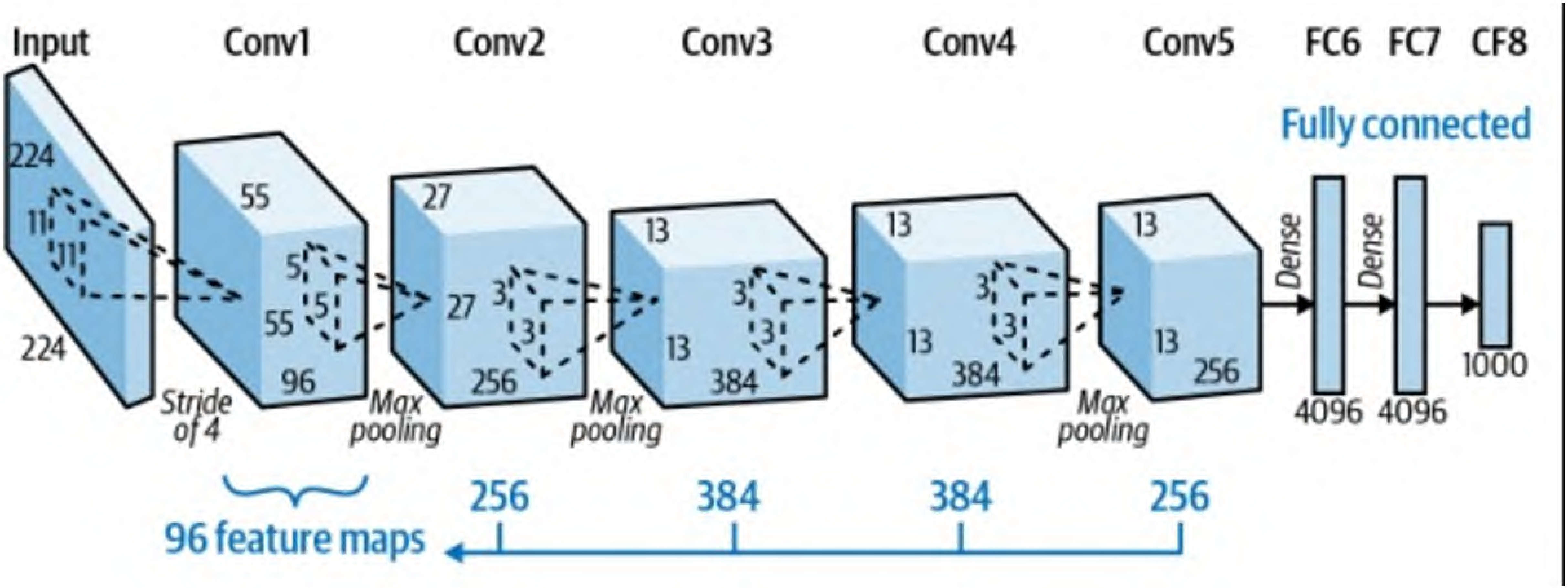
15用于图像分类的卷积神经网络

- 卷积神经网络（Convolutional Neural Network，CNN）是一种常用于图像分类的人工神经网络。它模仿生物视觉系统的工作原理，通过多层卷积和池化操作提取图像的特征，然后通过全连接层进行分类。在图像分类任务中，CNN已经在许多竞赛和实际应用中取得了很好的效果，如ImageNet比赛中的AlexNet、VGG、ResNet等模型。
- 此外，还有许多预训练好的CNN模型可供使用，例如TensorFlow Hub、PyTorch Hub等平台提供了许多经过训练的模型，可以直接应用于各种图像分类任务中。

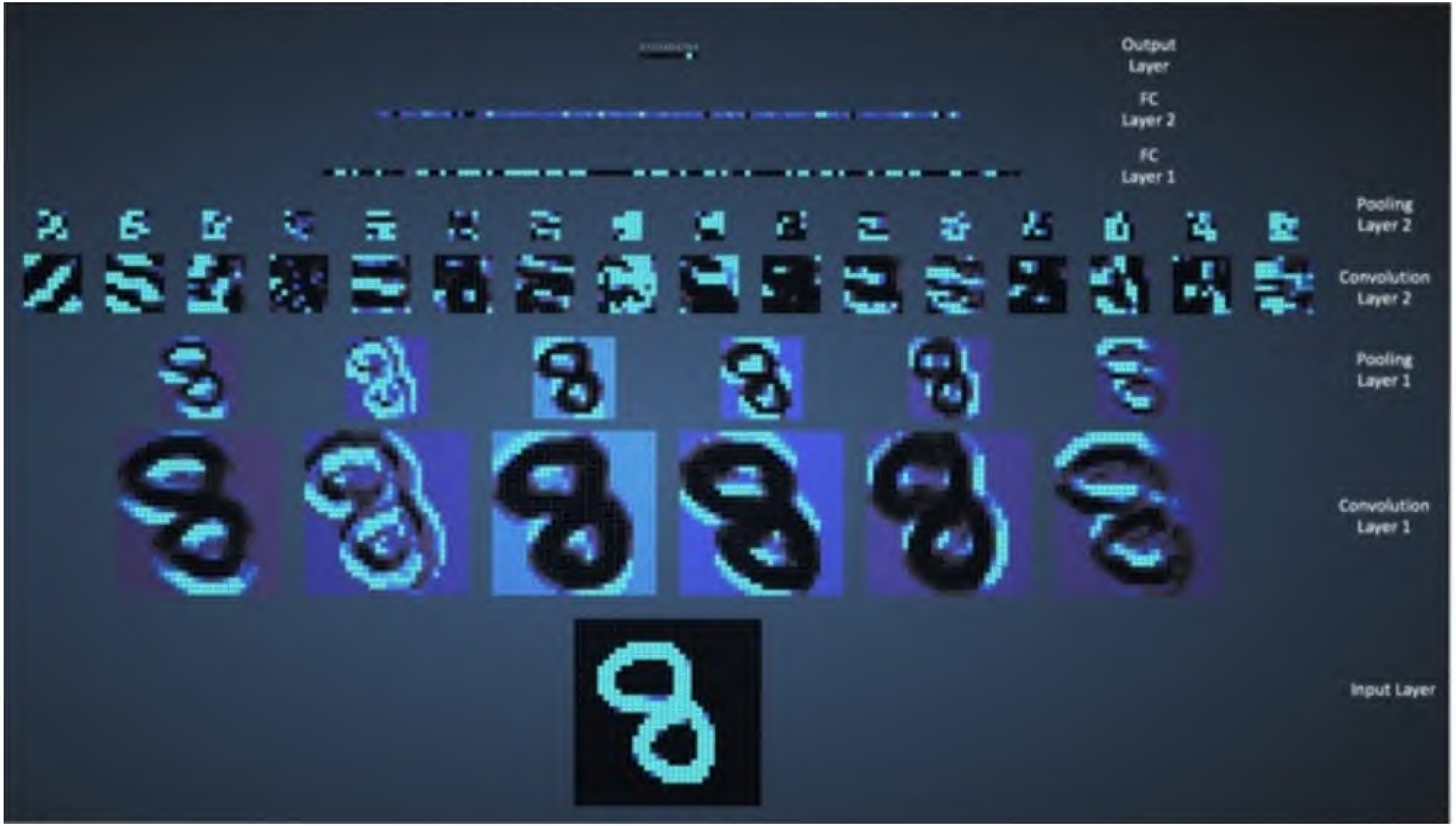
15用于图像分类的卷积神经网络



15用于图像分类的卷积神经网络



15用于图像分类的卷积神经网络



谢谢！

gulp@mail.las.ac.cn