

基于人眼视觉皮层系统的 SAR 图像目标识别

倪嘉成, 许悦雷, 马时平, 李 帅

(空军工程大学 航空航天工程学院, 西安 710038)

摘要: 人眼的视觉系统是已知最复杂最精确的目标识别系统; 针对现有浅层学习模型提取 SAR 图像目标特征效果不理想、普适性不强及目标识别精度不高等问题, 参考人眼视觉系统的分层特征提取机制提出了一种深层特征提取目标识别算法; 算法首先提出一种针对 SAR 图像的预处理方法, 在特征提取阶段应用多层稀疏自动编码器, 通过逐层减少神经元数目使编码器学习得到输入数据的压缩表示, 从而提取得到更抽象更鲁棒的目标特征; 采用 MSTAR 数据库 3 个目标进行实验, 试验结果表明, 算法能很好地对 MSTAR 目标进行特征提取和识别, 类比原有的方法有较高的识别准确率, 达到 96.7%。具有较强的实用性。

关键词: SAR 图像; 目标识别; 视觉皮层系统; 交叉视觉皮质模型; 稀疏自动编码器

SAR Target Recognition Based on Human Visual Cortical System

Ni Jiacheng, Xu Yuelei, Ma Shiping, Li Shuai

(Institute of Aeronautics and Astronautics Engineering, Air Force Engineering University, Xi'an 710038, China)

Abstract: A method for SAR image target recognition based on human visual cortical system is presented in this paper since human vision system is the most complex and accurate recognition system. First a new SAR Image preprocessing method is proposed. Then a sparse autoencoder is used to extract useful features by limit the number of cells in each layer. Experiment result of MSTAR public data show a better performance of recognition.

Key words: SAR; object recognition; visual cortex; ICM; sparse autoencoder

0 引言

合成孔径雷达 (Synthetic Aperture Radar, 简称 SAR) 是一种新型的微波成像雷达。由于具有高分辨率、全天候和全天时的成像特点, 以及大面积的数据获取能力, 对传统光学传感器起到了巨大的弥补和辅助作用, 能承担高科技军事战争中及时、准确地获取战场军事行动信息的重大任务, 因此对 SAR 图像的准确解译越来越受到人们的重视。对 SAR 图像中的目标进行识别是 SAR 图像解译的重要研究内容, 由于 SAR 图像中目标的分辨率不高, 细节特征不明显且含有大量相干斑噪声, 导致现有的目标识别方法识别率不高, 鲁棒性不好, 存在很大局限性。寻找高精度、高效和高鲁棒性的目标识别算法仍是基于 SAR 的自动目标识别 (ATR) 中的核心环节。

在一个 SAR-ATR 系统中, 最重要的环节就是找到更好的特征来表示图像中的目标, 文献 [1] 通过离散小波来提取特征, 并在识别阶段使用 SVM 进行分类。文献 [2] 使用主成分分析 (PCA) 来提取特征, 使用 ATR2 神经网络进行分类。文献 [3] 分别使用 PCA, LDE, NMF 来提取特征, 并使用决策融合的策略进行特征融合从而提高识别精度。文献 [4] 使

用散射中心的方法进行特征匹配。文献 [5] 使用目标的阴影部分进行特征提取试图找到更好的目标表示方法。文献 [6] 则使用稀疏表示的方法来提取特征。然而, 这些方法均忽略了一个事实, 那就是人眼视觉系统才是已知最好的特征提取系统。

人类, 甚至是最年幼的孩子对于图像的识别也没有任何的问题, 在发育到 25 个月时, 婴儿就已经具有识别三维目标的能力^[7]。由于人眼视觉系统具有优异的性能, 模仿人眼的视觉系统就成为计算机视觉的研究热点, 其中由深度学习算法构成的多层神经网络已经在图像识别领域取得了重大的成功。

深度学习其动机在于建立、模拟人脑进行分析学习的神经网络, 它模仿人脑的视觉皮层来解释数据^[8]。人眼视觉皮层系统的结构如图 1 所示。人眼视觉系统由多个皮层复合构成, 基本通路为 retina (视网膜) - LGN 层 - V1 层 - V2 层 - V4 层 - IT 层 (由 AIT 和 PIT 组成)。人脑视觉机制虽然分层, 但每层的神经元细胞并没有本质的不同, 只是因为所在的层级不同导致功能不同。深度学习正是应用了这种分层并且神经元细胞基本相同的特点, 构造结构基本相同的多层神经网络提取特征, 由 Hinton 等人于 2006 年提出后在模式识别领域取得了巨大成功^[9-10]。

本文将这种模拟人脑视觉皮层特性的多层提取特征的方法应用于 SAR 图像的目标识别当中, 提出了一种基于视觉皮层系统的 SAR 图像特征提取和目标识别方法。首先根据 SAR 图像的自身特性对图像进行预处理, 得到有利于学习的图像数据, 然后训练一个多层稀疏自动编码器 (sparse autoencoder) 对 SAR 图像进行特征提取, 通过 3 个隐藏层逐层提取特征, 从而得到更加抽象更加本质的特征表示。学习完毕后在分类阶

收稿日期: 2013-10-23; 修回日期: 2013-12-27。

基金项目: 国家自然科学基金 (61372167, 61379104); 航空科学基金 (20115896022)。

作者简介: 倪嘉成 (1990-), 男, 江苏南京人, 硕士研究生, 主要从事图像处理、机器学习及视觉认知方向的研究。

许悦雷 (1975-), 男, 河北辛集人, 副教授, 主要从事图像压缩、压缩感知、模式识别、目标检测与跟踪方向的研究。

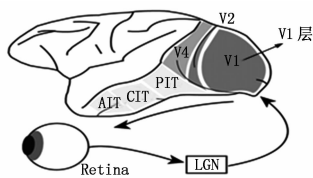


图 1 人眼视觉皮层结构图

段运用 softmax regression 进行 SAR 图像识别。

1 深度学习模型

本文使用的深度学习模型基于深度自动编码器, 由图 2 所示。深度自动编码器目的是提取输入目标的特征信息, 本文训练的多层自动编码器编码部分隐藏层设为三层, X 为输入图像像素数即维数, 可看作人脑的视网膜层。 n_1, n_2, n_3 分别为隐藏层的神经元数量, 每层代表视觉皮层中的一层, 且神经元数量逐层递减。由于限制了神经元数量, 迫使深度自动编码器学习输入数据的压缩表示, 从 X 维的向量中重构出 n_3 维。这样每层得到的参数都是输入 X 的一种表示, 从而提取了 X 的特征, 这样就完整构建了人脑的视觉皮层模型。

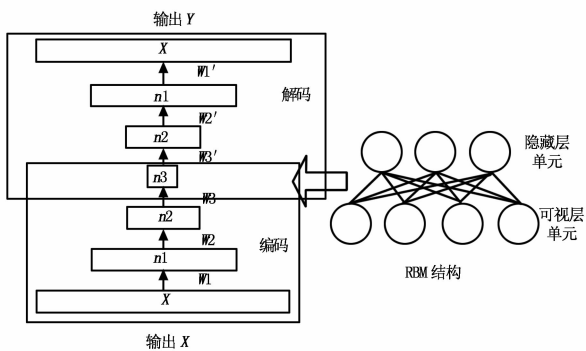


图 2 多层自动编码器和 RBM 结构图

1.1 受限波尔兹曼机

本文算法的深度自动编码器层与层之间由受限波尔兹曼机 (RBM) 构成。RBM 的训练过程如下:

类似于热力学, RBM 的能量定义为:

$$E(v, h) = - \sum_{ij} w_{ij} v_i h_j - \sum_i b_i v_i - \sum_j a_j h_j \quad (1)$$

其中: v_i 表示第 i 个可见单元的状态, h_j 表示第 j 个隐藏单元的状态, 我们假设所有的可见单元与隐藏单元均为二值, 即 $\forall i, j, v_i \in \{0, 1\}, h_j \in \{0, 1\}$ 。 w_{ij} 为连接权重, b_i, a_j 表示偏置 (Biases)。通过在可见单元与隐单元中各自增加一个状态为 1 的单元, 就可以通过连接权重表示偏置。故下文中略去了偏置。

利用能量函数和 Gibbs 分布, 我们可以得到关于 (v, h) 的联合概率分布:

$$P(v, h) = \frac{1}{Z} \exp(-E(v, h)) \quad (2)$$

$$Z = \sum_{v, h} \exp(-E(v, h)) \quad (3)$$

其中: Z 为剖分函数, 也可称为归一化函数。

由 RBM 的性质, 当给定可见单元的状态时, 隐藏单元的

激活概率是条件独立的, 记第 j 个隐单元的激活概率为 $P(h_j = 1 | v)$, 当 $h_j \in \{0, 1\}$ 时

$$P(h_j = 1 | v) = \text{sigm}(v^T w_{\cdot, j}) = \frac{1}{1 + \exp(-v^T w_{\cdot, j})} \quad (4)$$

其中: $w_{\cdot, j}$ 表示连接矩阵 W 的第 j 列。

同样地, 当给定隐藏单元的状态时, 可见单元的激活概率是条件独立的, 记第 i 个可见单元的激活概率为 $P(v_i = 1 | h)$ 。当 $v_i \in \{0, 1\}$ 时,

$$P(v_i = 1 | h) = \text{sigm}(w_{i, \cdot} \cdot h) = \frac{1}{1 + \exp(-w_{i, \cdot} \cdot h)} \quad (5)$$

其中: $w_{i, \cdot}$ 表示连接矩阵 W 的第 i 行。

由条件独立可以求得关于 (v, h) 的条件概率:

$$P(h | v) = \prod_j P(h_j | v) \quad (6)$$

$$P(v | h) = \prod_i P(v_i | h) \quad (7)$$

现在我们可以计算由 RBM 所定义的关于观测数据 v 的分布 $P(v)$:

$$P(v) = \sum_h P(v | h) P(h) \quad (8)$$

计算出观测数据 v 的概率分布 $p(v)$ 后, RBM 的参数 w_{ij} 可通过最大化 RBM 在训练集上的对数似然度学习得到:

$$\frac{\partial \log P(v)}{\partial w_{ij}} = \rho_{ij}^{(0)} - \rho_{ij}^{(\infty)} \quad (9)$$

其中: $\rho_{ij}^{(0)}$ 和 $\rho_{ij}^{(\infty)}$ 分别表示神经元 i 和 j 在零时刻和无穷远时刻的期望值。 w_{ij} 的变化量:

$$\Delta w_{ij} = \eta \left(\frac{\partial \log P(v)}{\partial w_{ij}} \right) = \eta (\rho_{ij}^{(0)} - \rho_{ij}^{(\infty)}) \quad (10)$$

其中: η 为学习效率参数。

1.2 稀疏自动编码器

传统的深度自动编码器在遇到数据量较大, 维数较高的输入数据时训练会变得很慢, 解决问题的办法是在隐藏层中加入稀疏约束, 这样就构成了稀疏自动编码器。加入稀疏限制后, 自动编码器在数据量很大时仍可以有效地学习特征。为了满足稀疏性的约束条件, 隐藏层节点输出的平均值应尽量为 0, 这样的话, 大部分的隐藏层节点都处于非激活状态。此时稀疏自动编码器的损失函数应为:

$$J_{\text{sparse}}(W, b) = J(W, b) + \beta \sum_{j=1}^{s_2} \text{KL}(\rho_j \| \rho_j^s) \quad (11)$$

后面一项为 KL 距离, 表达式为:

$$\text{KL}(\rho_j \| \rho_j^s) = \rho_j \log \frac{\rho_j}{\rho_j^s} + (1 - \rho_j) \log \frac{1 - \rho_j}{1 - \rho_j^s} \quad (12)$$

ρ_j 为稀疏性参数, 一般取接近 0 的数, 隐藏层节点输出的平均值为:

$$\rho_j = \frac{1}{m} \sum_{i=1}^m [a_j^{(2)}(x^{(i)})] \quad (13)$$

2 SAR 图像目标识别

提出了一种基于人眼视觉皮层系统的 SAR 图像目标识别算法, 识别算法流程如图 3 所示。

2.1 图像预处理

美国 MSTAR 计划录取的 SAR 地面静止目标图像包含 3 个区域: 目标 (Target)、阴影 (Shadow) 和背景杂波 (Back-

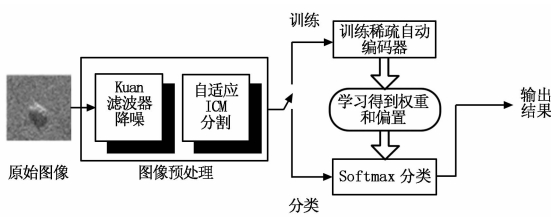


图 3 算法流程图

ground Clutters)。大量背景杂波及相干斑的存在使得视觉皮层系统的方法很难直接对目标进行识别。因此，本文提出了一种针对视觉皮层系统的 SAR 图像预处理方法。算法描述如下：

步骤 1：使用 Kuan 滤波器^[12]进行 SAR 图像降斑处理。Kuan 滤波器是 SAR 图像处理中经典的降斑方法，降斑效果较好，计算量较少。

步骤 2：使用自适应交叉视觉皮质模型（ICM，intersecting cortical model）对降斑后的图像进行目标分割。分割的目的是进一步去除背景杂波并减少数据量。ICM 具有生物神经网络中的信息传递延迟性和非线性耦合调制特性，相对于传统的人工神经网络模型更加接近实际的生物视觉神经网络。更重要的是，ICM 还模拟了视神经系统感受野受到刺激时，相邻连接神经元会同步激发 35~70 Hz 的振荡脉冲串^[13]。利用这种神经脉冲同步机制不仅能够做到高效且较为准确的分割，并且分割后图像更加符合人眼视觉特性，使下一步的特征提取变得简单高效。本文使用的自适应 ICM 算法能自适应的更改 ICM 中神经元的点火门限，增加了 ICM 算法的分割准确度。实验结果证明了该分割方法的有效性。

步骤 3：对处理过后的图像进行归一化，将灰度归一到 [0, 1] 范围中并通过图像剪切将目标至于图像的中心位置。目的是进一步方便稀疏自动编码器学习特征。

2.2 稀疏自动编码器提取特征

将预处理过后的图像通过稀疏自动编码器提取特征，具体步骤如下：

特征提取：

1) 输入：训练数据 X ，编码器层数 N ，训练周期数 E ；

权重和偏置的学习速率（隐藏层和可见层）： $e(w)$ ， $e(h)$ ， $e(v)$ ；

初始化参数 $\theta(w, a, b)$ ；

2) 输出：最优化参数 θ' ：权重 w' ，偏置 a' ， b' ；

步骤：

(1) 通过公式 (1)~(10) 计算输入层和第一隐藏层之间的参数 $\theta(0)$ 。

(2) 通过公式 (11)~(13) 计算损失函数 $J(\theta)$ 。通过梯度下降法迭代使损失函数不断减小从而优化参数 $\theta(0)$ 。

(3) 将第一隐藏层的输出作为与第二隐藏层之间 RBM 的输入，计算参数 $\theta(1)$ ，重复步骤 1, 2 直到求得第 N 隐藏层的参数 $\theta(N-1)$ 。

(4) 非监督学习完成，通过 BP 算法监督的自顶向下调节并逐层更新参数 θ' 。

(5) 训练完成，输出逐层提取的权重参数作为最终提取的特征。

训练完成后，将学习得到的权重参数代入 softmax regression 模型代替原有的权重参数，我们就可以对图像进行分类识别。

3 实验结果及分析

3.1 实验数据

本文使用的 SAR 图像是 MSTAR 项目组公布的 SAR 目标数据。文中选用其中 3 类目标，分别为 BMP2、BTR70 和 T72。分别选用 SAR 在俯视角为 17°和 15°时的数据作为目标的训练样本和测试样本。数据是 SAR 在方位角范围为 0~360°下的成像数据。

3.2 预处理

由于每个类别的图像数量相对过少，为了避免过拟合，本文提出一种通过灰度变换和微量缩放由单幅图像生成多幅图像的训练样本增加方法。将 1 幅图像通过灰度变换生成 6 幅图像，由微量缩放生成 4 幅图像，这样就将训练数据扩大了 10 倍，同时我们保持测试样本不变以便与其他算法比较，结果如图 4 所示。

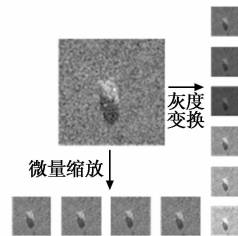


图 4 灰度变换和微量缩放的结果

增加训练样本以后进行图像预处理，图像预处理结果如图 5 所示，原图像素值为 128 * 128，剪切后的图像为 64 * 64。

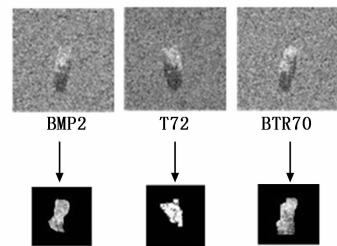


图 5 预处理前后的 3 种目标图像

表 1 不同输入数据时识别率对比 (%)

训练层数	训练样本数	预处理之前	预处理之后	
			Kuan	Kuan+ICM
一层	698	64.1	71.9	88.5
	6980	67.9	78.5	93.9
两层	698	65.0	73.9	90.5
	6980	68.8	80.8	95.4
三层	698	66.5	75.1	92.1
	6980	70.0	82.9	96.7

3.3 实验结果

通过对比不同隐藏层，不同输入训练样本和 SAR 图像预

处理前后的识别率来测试预处理算法的有效性和层数对算法的影响。表 1 列出了识别率对比情况。

由表 1 结果可以看出, 通过增加系统的隐藏层数确实可以提高系统对目标的识别率, 训练层数仅为 1 时, 预处理之前的识别效果较差, 仅为 64.1%。当增加隐藏层时识别效果有小幅度的提升, 当隐藏层为 3 时达到 66.5%。通过本文的样本增加方法和 SAR 图像预处理方法后算法的目标识别率有显著的提高。增加训练样本数后分类器效果有部分提升, 而当输入图像进行预处理之后识别效果显著提升, 最终当隐藏层数为 3 且经过预处理后算法的识别率达到 96.7%, 基本可以满足目标识别的需要。识别率的提升是因为增加训练样本数量可在一定程度上防止稀疏自动编码器出现过拟合现象。对图像的相干斑噪声抑制可以平滑图像, 减弱噪声对编码器中神经元的激活贡献率。自适应 ICM 分割则突出了编码器中神经元对目标部分的贡献率, 同时降低背景杂波的影响, 从而能够得到更好的识别效果。

下面通过与其他 SAR 图像目标识别方法的对比来检验本文算法的有效性。表 2 列出了识别率对比情况。

表 2 不同算法识别率对比

算法	平均识别率 (%)
PCA+SVM	82.14
参考文献 [2]	90.99
参考文献 [3]	92.51
参考文献 [4]	95.28
参考文献 [5]	96.1
本文算法	96.7

由表 2 可以看出, 本文算法的目标识别性能优于其余几种方法, 由于采用类似人脑视觉皮层的多层模型逐层提取特征, 本文算法较其余几种浅层学习模型能够更有效地提取目标特征。图 6 显示了将前两层隐藏层单元学习得到的权重可视化的结果。由图 6 可以明显看出, 第一隐藏层提取了目标的结构信息, 主要是各个方向的边缘, 角点特征。这些边缘特征对目标的表示和分类器最终的分类都有重要意义。提取完边缘特征后再向后传递到更深的隐藏层获得更加抽象的特征信息。这些都符合人脑视觉皮层对图像的处理方式, 对猫的初级视觉皮层(图 1 中的 V1 层)进行的试验(Hubel&Wiesel, 2000)表明, V1 层对边缘和角点信息比较敏感。

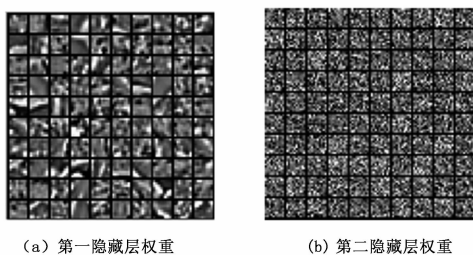


图 6 第一和第二隐藏层学习的权重可视化图

4 结束语

本文提出了一种基于人脑视觉皮层理论的 SAR 图像目标识别算法。该算法利用降斑算法和交叉视觉皮质模型对图像进行预处理, 消除相干斑噪声影响的同时突出目标信息, 方便进行特征提取。利用深层稀疏自动编码器模拟人脑视觉皮层分层的特点来提取特征, 利用 softmax 分类器实现目标的识别。实验结果表明本算法对 MSTAR 数据库中的 SAR 目标有很好的识别效果, 并且本文算法的优势还在于使用无监督训练, 适用于大规模输入数据环境下的应用场合。

参考文献:

- [1] 成功, 赵巍, 潘锦锋. 基于小波分解和支持向量机的 MSTAR SAR 目标分类识别研究 [J]. 中国图象图形学报, 2009, 14 (02): 317-322.
- [2] Ye X M, Gao W, Wang Y. Research on SAR images recognition based on ART2 neural network [A]. 7th IEEE Conference on Industrial Electronics and Applications (ICIEA) [C]. 2012.
- [3] Cao Z J, Cui Z Y, Fan Y, et al. SAR Automatic target recognition using a hierarchical multi feature fusion strategy [A]. GC12 Workshop: Radar and Sonar Networks [C]. 2012.
- [4] Tang T, Su Y. Object recognition based on feature matching of scattering centers in SAR imagery [A]. 5th International Congress on Image and Signal Processing (CISP 2012) [C]. 2012.
- [5] Papson S, Narayanan R M. Classification via the shadow region in SAR imagery [A]. IEEE Transaction on Aerospace and Electronic Systems [C]. 2012, 48 (2).
- [6] Thiagarajan J, Ramamurthy K, Knee P. Sparse representation for automatic target classification in SAR images [A]. Proc of the 4th International Symposium on Communications, Control and Signal Processing [C], 2010.
- [7] Wallis G, Bülthoff H. Earning to recognize objects [J]. Trends. Cogn. Sci, 1999.
- [8] Lee T, Mumford D. Hierarchical Bayesian inference in the visual cortex [J]. JOSAA, 2003.
- [9] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [A]. Advances in Neural Information Processing Systems (NIPS) [C]. 2012.
- [10] Ciresan D, Meier U, Schmidhuber J. Multi-column deep neural networks for image classification [A]. In IEEE Conf. CVPR [C]. 2012, 2: 3642-3649.
- [11] Le Roux N, Bengio Y. Representational power of restricted boltzmann machines and deep belief networks. Neural Computation [J]. 2008, 20 (6): 1631-1649.
- [12] 李涛, 许悦雷, 朱德晶. 基于边缘保持的 SAR 图像预处理方法 [J]. 计算机测量与控制, 2012, 20 (10): 1741-1746.
- [13] Ekblad U, Kinser J M. Theoretical Foundation of the Intersecting Cortical Model and Its Use for Detection of Aircraft, Cars and Nuclear Explosion Tests [J]. Signal Processing, 2004, 84 (7): 1131-1146.