

SpikeNet 的研究及其在快速人脸识别中的应用

张锦娟, 师 军, 于佳丽, 卢 照

(陕西师范大学 计算机科学学院, 陕西 西安 710062)

摘 要:对 SpikeNet 网络的研究,其主要目的在于根据电生理学的研究成果,通过计算机构造与人脑相似的神经元及其网络拓扑结构来模拟人脑的思维过程,进而学习人脑对外界事物的反应过程,实现对人脸图像中的兴趣区域识别和智能处理。对 SpikeNet 的研究将有助于提高网络的识别速度和识别效率,对人工智能的研究和发展具有深远的意义。SpikeNet 的理论基础是基于激发(spike)原理的排序编码(rank code ordering)研究思想。在数以万计的视觉神经元中,大脑可以将神经元按照敏感性排序,对于特定情况敏感的物体,特定的神经元会及早突破阈值而激发产生 Spike,并通过多层网络不断传递信息。对人脸识别的实验表明,SpikeNet 在速度上达到实时的神经计算,并且能很好地克服对比度和噪声的影响,获得比较理想的实验效果。

关键词:机器视觉;SpikeNet;排序编码;前馈式神经网络

中图分类号:TP183

文献标识码:A

文章编号:1673-629X(2010)07-0235-05

Research on SpikeNet and Its Application in Quick Face Recognition

ZHANG Jin-juan, SHI Jun, YU Jia-li, LU Zhao

(College of Computer Science, Shaanxi Normal University, Xi'an 710062, China)

Abstract: The significance of the research on SpikeNet is to construct the neural network which could simulate the signal processing in brain neural network through the computer. This could help us with image processing on learning how the brain reflects the environment in visual field. The research could help to improve both the processing speed and the rate of image recognition. The basic idea behind SpikeNet is rank code ordering that there are billions of neurons in the visual system, brain could use the order in which neurons fire spikes to code information with different sensitivity. For the sensitive neurons, they can fire a spike first which present the different order of spikes and with these spikes, network could pass the information on and on. The face testing with different contrast and luminance values shows that the result of the experiment indicates that our system is feasible.

Key words: machine vision; SpikeNet; rank order coding; feed-forward neural network

0 引言

随着自动人脸识别系统的提出,并且由于其在安全访问控制、视觉监测、基于内容和检索和新一代人机界面等领域的应用价值,使得人脸识别开始作为一个独立的课题受到研究者的普遍重视。

传统神经网络是一种应用较为广泛的人脸识别方法,但是不可避免地存在着以下问题:

第一,网络在工作时,需要每一个神经元都参与矩阵的计算,以致大量的神经计算降低了神经网络识别

的速度,不符合生物视觉反应的要求;

第二,算法的学习速度很慢;

第三,出现所谓“过拟合”现象^[1],不能反映样本内含的规律,因此难以得到较好的识别结果。

基于 SpikeNet 神经网络的人脸识别方法,它是对数以万计的视觉神经元按照敏感性排序、特定的神经元激发来表征信息,无需所有神经元参与运算,减少计算量,提高速度;同时降低对样本需求的数量,容易提取样本中的相同趋势,大大提高了识别效率。文中采用基于 SpikeNet 神经网络实现人脸识别,其主要意义在于根据电生理学的研究成果,通过计算机构造与人脑相似的神经元及其网络拓扑结构来模拟人脑的思维过程^[2,3],进而学习人脑对外界事物的反应过程,实现对图像中的兴趣区域识别和智能处理。从而网络的识别速度和识别效率大大提高。

收稿日期:2009-09-01;修回日期:2009-12-01

基金项目:陕西省自然科学基金资助项目(2005A12)

作者简介:张锦娟(1983-),女,陕西西安人,硕士研究生,研究方向为人工智能信息处理;师 军,副教授,硕士研究生导师,研究方向为人工智能信息处理、并行计算等。

1 SpikeNet 网络算法设计

1.1 SpikeNet 神经网络与传统网络的比较

传统的神经网络主要是将特征矩阵与权值相乘,然后根据相乘结果的不同来确定对事物的分类。这种传统的网络工作时,需要每一个神经元都参与矩阵的计算,每一个神经元计算的结果都会对最终分类产生影响,以致大量的神经计算降低了神经网络识别的速度,不符合生物视觉反应的要求。

传统的人工神经元设计也相对简单,它们只是简单地用一个在 0~1(或 -1~1)之间的浮点数来代表神经元参与计算。很明显,这种设计方法没有很好地表现出神经元细胞的结构,也不能实现达到神经元对信号分类的功能。

具体可以概括为以下的缺陷:算法的学习速度很慢;网络训练失败的可能性较大;难以解决应用问题的实例规模和网络规模间的矛盾。这涉及到网络容量的可能性与可行性的关系问题,即学习复杂性问题;网络结构的选择尚无一种统一而完整的理论指导,一般只能由经验选定。新加入的样本要影响已学习成功的网络,而且刻画每个输入样本的特征的数目也必须相同;网络的预测能力(也称泛化能力、推广能力)与训练能力(也称逼近能力、学习能力)相矛盾。一般情况下,训练能力差时,预测能力也差,并且一定程度上,随训练能力的提高,预测能力也提高。但这种趋势有一个极限,当达到此极限时,随训练能力的提高,预测能力反而下降,即出现所谓“过拟合”现象。此时,网络学习了过多的样本细节,而不能反映样本内含的规律。

对于这些问题的来源,主要的原因来自于传统网络的拓扑结构和基本的神经元的设计上。传统的设计方法一般主要是基于同步的神经元作为神经网络的输入,通过权值的直接乘积、求和来对中间信号做处理。

而在异步网络中,采用激发型的神经元来构建神经网络,并应用激发信号来表征计算后的神经元的编码信息,这样做在很大程度上可以提高神经网络的处理的性能,减少不必要的计算量,加快了图像识别的速度。SpikeNet 即是根据此激发式神经元构建的一个简单的前馈式网络。采用此网络的设计方法可以很好地克服上述一些问题。最大的优势,它可以极大地提高网络计算和识别的速度,同时可以降低对样本需求的数量,对于各个方向的样本,容易提取样本中的相同趋势,并可以找到对应的权值来表现这样一种趋势,在识别过程中更容易达到提高识别率的效果。

1.2 SpikeNET 网络中神经元的设计

激发式神经元的原始模型与传统神经元大体上相同,其中神经元设计时包括电位值(membrane poten-

tial)、阈值(threshold)、电位时间常量(membrane time constant)。具体而言,对于两个神经元 A, B, 权值 weight, 其中 A 为第一层的输入, B 为下一层的输出。它们之间的关系为: $B = A * \text{weight}$, 其中, 在数学表达中, A 和 B 分别代表两个神经元的电位值。

网络中每一个神经元的计算由第一层的输入与权值的乘积表示, 神经元的电位值为: $A = \sum w_i e_i$, 其中, A 即为接受输入后的神经元的电位值。经过阈值函数 f 后, 再判定输出。与传统网络不同的是, $f(A)$ 并不是直接的输出, 而是根据该层激发的顺序相应输出。具体操作如下所示:

1) 接受输入;

2) 判断 A 是否超过阈值;

3) 若超过, 则激发一个 spike, 并将该 spike 存入 spike-list 中^[4], 并将 A 与阈值的差值存放在原神经元中作为电位值存放;

4) 若未超过, 则保存 A 值为神经元电位值。继续等待, 并累加后续到来的激发。直到时间段消失, 其间若 A 超过, 则执行步骤 3), 若为超出, 则继续等待。

在时间段(time-stop)内, 对于添加到 spike-list 中的激发神经元, 必定有激发的先后顺序, 对此排序编码(rank order coding), 可以表征图像上相应的信息, 如最早激发的神经元, 更“像”识别目标的区域。而对于没有激发的神经元, 在后续的网络中将不参与计算。后续网络的计算和识别只会涉及到激发后保存在 spike-list 中的神经元, 这样做产生的作用可以大量减少参与计算的神经元数目。时间段(time-stop)是 spike 网络中的重要参量, 它的大小设定直接影响了激发的神经元数量, 对后续网络的处理产生一定影响。

容易推断, 当时间段(time-stop)的值设定较大的时候, 激发的神经元数量会相对比较多, 这是由于网络运行时神经元接受上一层输出累加的结果。而相应的, 当时间段(time-stop)的值设定较小的时候, 激发的神经元数量会相对比较少。

1.3 SpikeNet 排序编码(rank order coding)

排序编码是指按照神经元激发的先后顺序编码, 用不同的顺序来表征信息。如 A, B, C, D, E, F, G, H, I, J 这 10 个神经元在收到输入后累加, 设其累加和为: $C > B > D > A > E > F > G > J > H > I$, 由此顺序可以构成 10! 个排序方案, 即表征 $10! = 3628800$ 种信息。

在神经元方面可以用图 1 表示, 对于先激发的神经元(在上述 10 个神经元中, 按照趋势, 则 C 和 B 最有可能首先激发), 其激发的时间会相应较短。在一定时间内, 激发后的神经元按照先后顺序排序, 如图 1 所

示,最左侧 Spike 先于中间和右侧的 Spike,其激发时间最短,受到的影响最少。对于先激发的神经元,其激发的时间会相应较短。在一定时间内,激发后的神经元按照先后顺序排序,激发时间最短的,受到的影响最少。

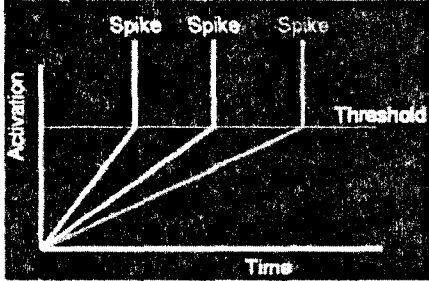


图 1 神经元在不同时间突破阈值激发 Spike 而形成有顺序的编码

对于排序后的神经元按照以下方式输出:

令 $A\{a_1, a_2, \dots, a_{m-1}, a_m\}$ 为输入的神经元, $W\{w_{1,i}, w_{2,i}, \dots, w_{m-1,i}, w_{m,i}\}$ 为神经元间的权值, mod 为 $0 \sim 1$ 的调节因子, Activation 为神经元的电位值。那么,当神经元工作时,对于神经元 i 在 t 时刻,其电位值为:

$$\text{Activation}(i, t) = \sum_{j \in \{1, m\}} \text{mod}^{\text{order}(a_j)} w_{j,i}$$

在上述公式中, $\text{order}(a_j)$ 是神经元 a_j 在神经元 A 中的激发顺序。当 $\text{Activation}(i, t)$ 大于阈值时,神经元 i 激发,并释放一个 spike 向下一层传递。如此反复计算。

有两点可以从此机制中推断:

1) 当特征的模型与权值大小的先后顺序越相一致时,表征目标的神经元越容易激发。尤其是当权值的顺序与样本的特征完全符合时,输入的神经元将会最容易提高其自身的累加电位值。

2) 最容易提高累加电位值的神经元最容易激发并释放 Spike。

这两点在神经网络对权值的训练中起到了最关键的作用。

1.4 SpikeNet 网络训练方法

权值的选取和使用在神经网络的识别中起着关键的作用。在异步的神经网络中,网络的一端是图像的各像素,另一端是识别的结果(仍用像素表示)。由于异步激发的随机性,矩阵的大小并不固定,这样就不采用其他的训练方法。目前, SpikeNet 的研究仅仅只用于设计前馈式网络,经常采用的训练方法是使用样本的中心点位置(即最终激发的神经元)方向推导每一层的网络权系数^[5~7]。

按步骤,学习过程如下:

1) 在网络学习开始前,将样本中需要识别的像素中心坐标(即最后激发的神经元)用手工找出并标识;

2) 使用该中心坐标推导前一层的与之相关的神元元的权值;

3) 依次类推,可以找到识别所需的每一层相关坐标下神经元对应的权值。

具体的做法如下:

对神经元 $A\{a_1, a_2, \dots, a_{m-1}, a_m\}$, 操作时使用下列公式计算各神经元间的权值:

$$\Delta W_{j,i} = \frac{\text{mod}^{\text{order}(a_j)}}{N} \quad (1)$$

$\Delta W_{j,i}$ 即是每一次有神经元激发时对权值的改变量,在 N 幅图像中,一共改变 N 次,最后取得的便是描述样本的“平均权值”。这种提取方法可以很好地找到各样本的共同的相似性,克服了传统神经网络的局限。

2 面向人脸识别的 SpikeNet 设计

在网络结构设计上,依据电生理学的原理,讲述图像处理的流程,一般情况下,视觉图像依次经过视网膜(V1)才到达识别层^[8](如图 2 所示)。本节通过对人脸定位与识别的实验来描述通过 SpikeNet 对视觉图像处理的流程。

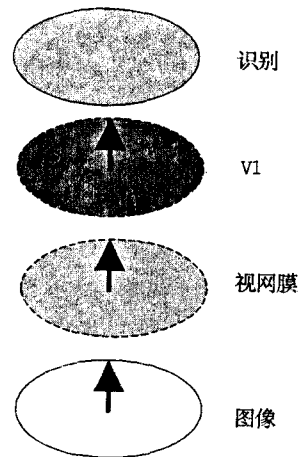


图 2 SpikeNet 中使用的网络设计结构

在 SpikeNet 网络平台之上,以人脸中心定位为例,搭建了一个如图 2 所示的软件系统,在神经网络的设计上,采用简单的四层前馈式神经网络^[9~12],其中第一层采用对比度映射层来模拟视网膜对视觉图像的处理,第二层采用方向映射模拟 V1 层对视网膜处理后的信息的处理,第三层、第四层分别采用特征映射和人脸定位映射来确定最后的识别中心。在权值的选取上,后面两层采用对样本图像抽样的方法来构造权值矩阵。网络的基本构架如图 3 所示。

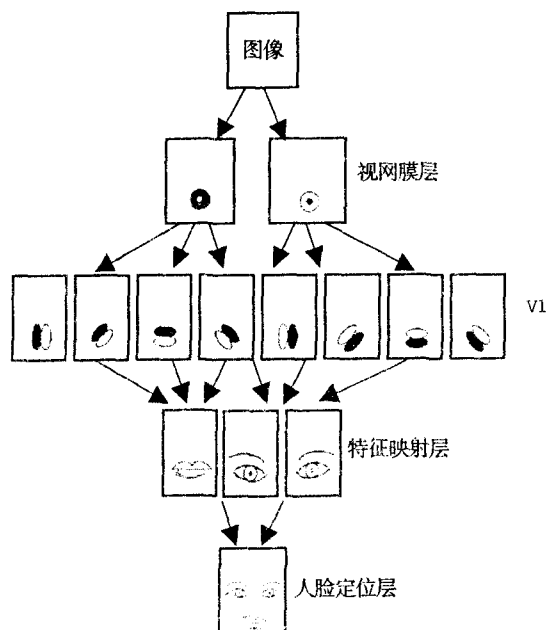
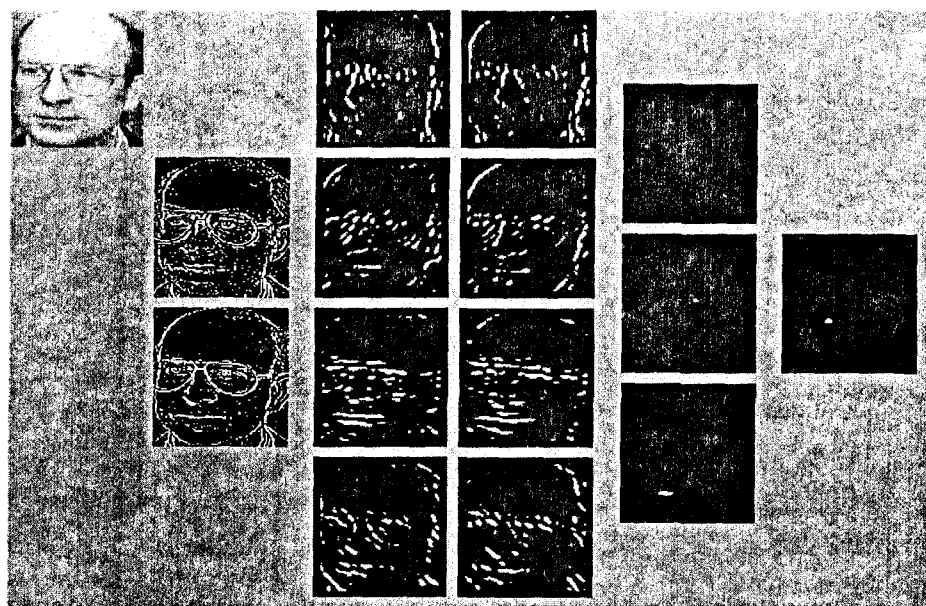


图3 人脸定位实验中网络设计的基本构成示意图

3 实验及结论

在本软件系统中,设计了对 40 张人脸的识别实验,训练过程由式(1)来获得。40 张人脸中,每张脸用 8 幅图片训练,最后识别时,每人使用 10 张图片。图像的处理过程如图 4 流程所示,按照前馈式网络的顺序相应处理。其中,输入的原始图像为 92×112 的灰度图像,图像处理的具体过程如下所示:



(第1列:原始图;第2列:突出边缘;第3、4列:8个方向上的边缘信息;
第5列:分别在左眼、右眼和嘴巴处激发;第6列:人脸中心定位)

图4 图像处理的具体过程

第一层:对比度映射层。这一层使用了简单的锐化算子对图像做卷积,可以滤除图像中对比度信息较

弱的部分,从而突出了人脸中的边缘。使用的卷积算子为:

$$\begin{matrix} -1 & -2 & -1 & 1 & 2 & 1 \\ -2 & 12 & -2 & 2 & -12 & 1 \\ -1 & -2 & -1 & 1 & 2 & 1 \end{matrix}$$

其中对比度算子可以有不同选取方案,可以依照图像的拍摄环境和拍摄效果,而其目的是模拟视网膜层(Retina)对视场图像对比度的分离。在图3中,每一张人脸图片首先经过视网膜层(分为 ON 和 OFF),得到不同的对比度图,然后 8 个方向求边缘,基本可以获得完整的人脸边缘图像。在与样本中各人脸中心的对应训练中(其中,人脸中心在此处被认为是左眼、右眼、嘴巴三点构成的三角形的中心,且此中心由操作者手工获得并写入文件中),不断通过 $\Delta W_{j,i}$ 改变 $W_{j,i}$,达到训练的效果。

第二层:方向映射层。这一层使用 8 个方向算子对锐化后的图像分别求其方向上的边缘信息,使其在 8 个方向上成图。其结果是该层网络上的神经元对于边缘信息丰富的部位更加敏感,提前激发,从而滤除了人脸图像中的无边缘信息区域。

第三层:特征映射。神经元与网络权值作用,分别在眼睛、嘴巴处神经元激发。

第四层:人脸定位映射。根据眼睛和嘴巴的位置,使用高斯滤波器求取中心,该中心即为人脸的中心。

识别过程即为图像与权值作用的过程。首先将像

素值转变为用数据结构编写的链表形式,称为 spike-list,通过对比度映射层,得到两个对比图,然后经过边缘算子提取各方向边缘。图像经过 8 个方向的边缘卷积后,提取出 8 个边缘图,对这 8 个图分别对其训练好的权值相乘,按照激发-排序-编码的过程获取识别结果。最后的识别如图 5 所示。可以清晰地看到,输入样本人脸,网络会即时地作出反应识别。对识别出的人脸会在图像后的黑色区域显示一个白色的亮点。神经网络的权值对应其他人脸的神经元均未激发。正如权值训练后的排序一样,神经元被很好地按照图片的“相似性”归类,从而识别特征。

识别过程即为图像与权值作用的过程。首先将像

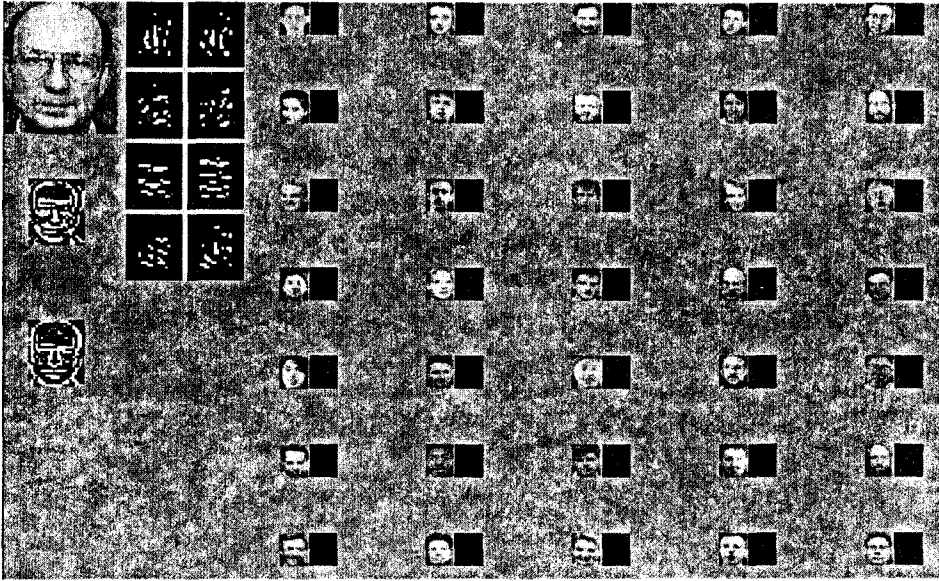


图 5 人脸识别结果

在 Linux 下 KDevelop 平台之上,对 400 张人脸图(公开人脸数据库)做了识别测试,首先将这 400 张图像放入网络训练,然后将生成的权值矩阵替换原始的权值,将这 400 张图像放入网络识别。实验表明,对于训练后的图像,网络均能按照要求对应识别,此 400 张人脸图识别中,识别结果的成功率为 100%。

4 结束语

随着研究的深入,传统神经计算的局限性已经变得明显起来。分析原因,全并行计算和固定的连接结构是根源所在。SpikeNet 作为目前最新的神经网络成果,克服了传统神经网络模式带来的不便性。总的来说,从传统网络到 SpikeNet 是一项创新的科研成果。但是,目前的 SpikeNet 还不能很好地模拟快速视觉的信号处理,还有很多的问题有待解决。

参考文献:

[1] 李 伟. 人脸识别算法在智能手机上的实现[J]. 计算机技

术与发展,2008,18(1):161-163.

[2] 沈 虹. 基于 Spike 神经元的神经网络及其股市预测模型[J]. 商场现代化,2008(12):248-249.

[3] 衣文文,杨彬彬,胡彦磊,等. 一种基于外形区域的图像配准方法与实现[J]. 计算机技术与发展,2008,18(4):1-4.

[4] 石争浩. 神经网络图像匹配技术研究[D]. 西安:西安微电子技术研究所,2004.

[5] 吴健康. 神经网络图像识别[R]. 合肥:中国科学技术大学,1990.

[6] 李启娟,李金屏. 基于轮廓信息的人脸检测[J]. 计算机技术与发展,2008,18(9):108-111.

[7] 孙晓玲,侯德文,储凡静. 人脸识别中的眼睛定位方法[J]. 计算机技术与发展,2008,18(10):46-48.

[8] Thorpe S, Delorme A, VanRullen R. Spike - Based Strategies for Rapid Processing[J]. Neural Networks, 2001, 14(6-7):715-725.

[9] VanRullen R, Gautrais J, Delorme A, et al. Face processing using one spike per neurone[J]. Biosystems, 1998, 48(1-3):229-239.

[10] Delorme A, Thorpe S. SpikeNET: an event - driven simulation package for modelling large networks of spiking neurons[J]. Network, 2003, 14:613-627.

[11] Oram M W, Perrett D I. Time course of neural responses discriminating different views of the face and head[J]. Neurophysiol, 1992, 68:70-84.

[12] Thorpe S J, Gautrais J. Rank Order Coding[C]//Computational Neuroscience: Trends in Research 1998, New York: Plenum Press, 1998:113-118.

(上接第 234 页)

球物理学进展,2007,22(5):1466-1473.

[7] 雷英杰,张善文,李续武,等. MATLAB 遗传算法工具箱及应用[M]. 西安:西安电子科技大学出版社,2005:11-13.

[8] 张跃飞. 基于稀疏分解的图像压缩[D]. 成都:西南交通大学,2006.

[9] 陈发宇,杨长春. 基于 MP 方法的地震信号快速分解算法[J]. 地球物理学进展,2007,22(6):1692-1697.

[10] 马国田,梁昌洪. 基于混合遗传算法的宽带阶梯阻抗变换

器的优化设计[J]. 西安电子科技大学学报,1999,26(1):8-12.

[11] 陈学华. 时频分布与地震信号谱分析研究[D]. 成都:成都理工大学,2006.

[12] Castagna J P, Sun S, Siegfried R W. Instantaneous spectral analysis: Detection of low - frequency shadows associated with hydrocarbons[J]. The Leading Edge, 2003, 22:120-127.