DOI:10.19651/j.cnki.emt.2106562

# 多种构图方式下的加密流量分类

朱文斌 马秀丽
(上海大学 通信与信息工程学院 上海 200444)

**摘 要**：传统网络流量分类方法难以区分使用 VPN 加密的网络流量，为了实现加密流量的分类，提出了一种基于多种构图方式的网络流量图像分类方法。研究了 5 种特殊的构图方式，将加密网络流量转换为流量图像，最后利用卷积神经网络进行分类。通过在自己采集的 VPN 加密流量数据集和 ISCX VPN-nonVPN 公开数据集上的实验结果表明，此加密流量分类方法的分类精度在两个数据集上分别达到 90%以上与 95%以上。对角型或瀑布型构图方式的分类精度较传统线型构图方式有 1%左右的提升。特殊的构图方式能加强流量图像中像素点的相关性，增加流量图像中的图像特征，实现流量分类精度的提升。

**关键词**：加密流量分类；构图方式；卷积神经网络
**中图分类号**：TP399 **文献标识码**：A **国家标准学科分类代码**：510.50

# Classification of encrypted traffic based on multiple composition methods

Zhu Wenbin Ma Xiuli
(School of Communication and Information Engineering, Shanghai University, Shanghai 200444, China)

**Abstract**: Traditional network traffic classification methods are difficult to distinguish network traffic encrypted using VPN. In order to achieve classification of encrypted traffic, a network traffic image classification method based on multiple composition methods is proposed. Five special composition methods are studied to convert encrypted network traffic into traffic images, and finally convolutional neural networks are used to classify them. The experimental results on self-collected VPN encrypted traffic dataset and ISCX VPN-nonVPN public dataset show that the classification accuracy of this encrypted traffic classification method reaches more than 90% and 95% on the two datasets, respectively. The classification accuracy of diagonal or waterfall composition method has about 1% improvement over the traditional line composition method. The special composition method achieves the improvement of encrypted traffic classification accuracy by enhancing the correlation of pixel points in the traffic image and increasing the image features in the traffic image.

**Keywords**: encrypted traffic classification; composition method; convolutional neural networks

## 0 引 言

近年来，网络流量分类是一个十分热门的研究领域。流量分类是根据特定要求将相关的网络流量划分为多个优先级或多个服务类。其目的就是加强对网络的管理，其在网络安全、网络任务管理以及网络规划等方面都有着十分重要的作用[1]。各种各式 VPN 的出现，为流量分类带来了巨大的挑战。这些 VPN 使用特定的加密技术对网络流量进行加密，来躲过防火墙和网络入侵检测系统的检测[2]。现在市面上 VPN 的种类繁多，不同 VPN 使用的加密技术不同。除此之外，同一种 VPN 也会在一段时间内定期进行更新，改进加密方式，这些特点都会使得一些传统的流量分类方法失去作用[3]。

近十几年来，机器学习得到了快速的发展，并在许多领域取得了成功，其中深度学习最受研究人员的青睐。因此，近些年来，部分研究者尝试将网络流量转换成图像，利用卷积神经网络在图像识别的成功，来实现对网络流量的分类[4]。Wang 等[5]提出了一种端到端的加密流量分类方法，这是一种不需要流量统计特征的方法，直接将原始流量作为一维卷积神经网络的输入数据，实现流量分类。Lotfollahi 等[6]提出，直接利用每个 IP 数据包的有效内容作为网络的输入，并使用卷积神经网络和栈式自动编码两个算法对流量进行分类。Shapira 等[7]提出，利用数据包到达时间与数据包大小两种统计特征，将网络流量转换为图片，然后使用卷积神经网络来识别流量类别。

收稿日期：2021-04-30

上述研究者们，大多是使用线型排列方式将网络流量转换成图像。这种直接的线型排布会破坏数据包内部邻近数据间的相关性，从而减少流量图像的有效特征，影响分类结果[8-9]。在本文中，提出了多种新颖的构图方法，将加密网络流量转换成图像，利用深度学习算法对转换后的图像进行分类，进而实现对加密流量的分类。不同的构图方法，能加强流量图像中相邻像素点的相关性，提升网络流量分类的精度。为了验证所提出方法的可行性，分别对一个公共流量数据集 ISCX VPN-nonVPN 和人工手动采集的流量数据集进行实验。

## 1　流量分类的多种构图方式研究

### 1.1　流量图像

网络流量实际的表现形式就是字节流。如图 1 所示，长度为 784 Byte 的一维向量。将 784 Byte 的一维向量转换成 28×28 的二维矩阵，再将其转换成灰度图像，得到的图像就称为流量图像。这种将网络流量转换成流量图像的方法，目的是为了将卷积神经网络应用到流量分类当中来。但是不同于自然中真实的图像，流量图像只是利用了维度转换的方式将一维的字节流转换成二维矩阵。因此，考虑不同的构图方式，是本文的研究重点。

| byte1 | byte2 | byte3 | ... | byte784 |
|---|---|---|---|---|

图 1　一维字节流

### 1.2　多种构图方式

目前的研究基本都采用的是线型构图方式，将数据包转换为流量图像。当网络设备进行通信时，其发送或接受的数据包是在相关通信协议或是加密协议下产生的。因此，相邻数据包之间有一定的相关性，同一数据包的相邻字节间也有一定的相关性。图 2 是利用 16 Byte 长度的向量进行各种构图方式的结果。如图 2(a)所示在线型构图方式下，横向上相邻的字节是连续的。但在纵向上，紧靠在一起的字节在数据包中的实际位置有很大差别，并不是连续的。例如，byte1 连接着 byte5，byte7 连接着 byte3 和 byte11。在真实的图像中，各个像素并不是独立存在的，像素之间有很大的相关性。许多图像处理方法也都利用了像素之间的相关性，例如图像压缩、图像恢复。想要使流量图像更加地接近真实图像，就要考虑流量图像中像素点的相关性。因此，考虑将一维字节向量中连续的字节尽可能在二维图像的排布中相近。

为了增加流量图像种各像素点的相关性，设计了 5 种特殊的构图方式。如图 2(b)～(f)所示，对角型构图方式、瀑布型构图方式、倒流瀑布型构图方式、螺旋型构图方式、反螺旋型构图方式[10]。图中的箭头方向表示了字节流排布的顺序。对角型构图方式，依据对角线对一维字节流进行排布，从左上角开始排步，相邻的字节靠近在一起，例如，byte1 紧挨着 byte2 和 byte3。瀑布型构图方式是参考瀑布流动时的形态得出的构图方式，沿着对角线往横、纵两个方向排列，这也是一个新颖的尝试。反瀑布型构图方式，是依据瀑布型构图方式按相反的原理进行排布。螺旋型构图方式，这是一种由内螺旋向外的排列方式。在这种方式下，字节流的前半部分紧密排列在流量图像的中心，可以验证字节流前半部分之间的相关性强弱。反螺旋型构图方式，则与之前完全相反，是一种由外螺旋向内的排列方式。在这种方式下，字节流的后半部分紧密排列在流量图像的中心，可以验证字节流后半部分之间的相关性强弱。数据包的前半部分是包头，里面有一些比较重要统计特征，而后半部分是 payload，里面是传输数据。不同的螺旋方式可以判断出数据包头部字节间的相关性强，还是 payload 中字节的相关性强。

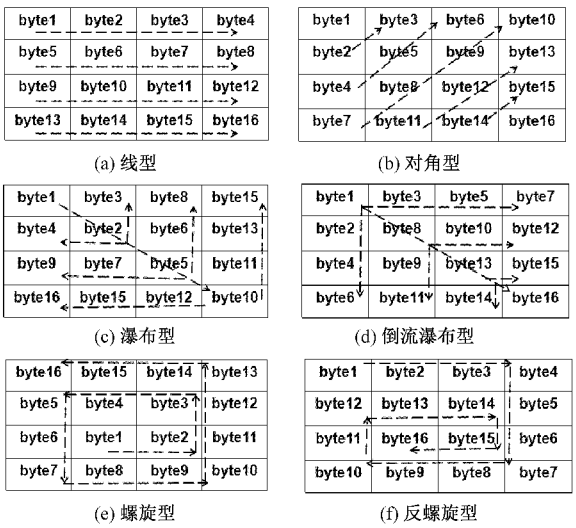


图 2　多种构图方式

## 2　基于卷积神经网络的流量图像分类

### 2.1　数据集的选择

目前大部分的研究都基于一些公开数据集，例如 ISCX 数据集。由于互联网技术的高速发展，一些公开数据集的流量数据可能不具有很强的时效性与代表性。但是，若只使用人工采集的私人流量又会影响结果的可信度[11]。因此，本文分别使用了公共数据集与人工捕获的流量数据进行实验。

1)ISCX VPN-nonVPN 数据集

Habibi 等[12]发布了此数据集，其是目前最热门的公共数据集之一。此数据集分为两大类，VPN 和 nonVPN。每一类中，分别包含 7 种不同的流量类型，一共 14 种流量类型。此数据集是根据会话期间进行的特定活动进行标记的，例如语音通话、聊天、文件传输或视频通话。除了此种标记方式之外，还选择了另外一种标记方式，对不同应用程序下的各种行为进行标记。此种方法总共标记了 22 类的

流量标签。

2)人工捕获的数据集

本实验的研究重点是考虑不同 VPN 下加密流量的分类。ISCX 数据集的流量数据只标记了 VPN 或是 nonVPN，无法获知他们使用了哪款 VPN 进行流量加密。随着互联网技术的发展，现在市面上有各式各样的 VPN 供人们使用，不同 VPN 使用的加密方式也不同。因此，只考虑一种 VPN 下加密流量的分类在实际应用中并没有意义。本文使用了目前最热门的几款 VPN 来产生加密流量，ExpressVPN、SsrVPN、PsiphonVPN、V2rayVPN。不同 VPN 下，使用了目前最热门的几款国内外 APP。表 1 所示为所捕获的数据集的详细内容。

**表 1 捕获的流量类型**

| VPN 种类 | 应用程序 |
|---|---|
| ExpressVPN | facebook telegram twitter youtube bilibili wangyiyun wechat |
| PsiphonVPN | facebook telegram twitter youtube bilibili wangyiyun wechat weibo |
| SsrVPN | facebook telegram twitter youtube bilibili wechat weibo |
| V2rayVPN | facebook telegram twitter youtube wechat |

### 2.2 数据预处理

数据预处理是十分关键的步骤。本文是在数据包级别上产生流量图像，再利用卷积神经网络网络进行流量图像分类。因此，对于数据包的筛选过程，以及数据包的预处理十分重要。在流量采集的过程中，由于后台其他应用程序的活跃或是 VPN 网络连接不稳定等情况下，会产生一些干扰的数据包，需要对这些无效数据包进行过滤。剩余的有效数据包需要进行以下预处理过程[13]。

数据预处理过程大致分为 4 步：

1)过滤无意义的会话。利用会话的五元组定义，分离出所需的会话。相同会话中数据包拥有相同的源 IP 地址、源端口、目的 IP 地址、目的端口和传输层协议。

2)过滤无意义的数据包。一个会话当中，包含两个地址间发送的所有数据包。其中有一些对流量分类没有意义的数据包，例如 ARP 和 ACK。这些数据包只是基于 UDP 或是 TCP 协议下建立会话链接而产生的包，并没有携带有效的数据内容。

3)截断或拼接。由于本实验基于经典的卷积神经网络模型，因此需要固定输入数据的长度 $n$ Byte(本文中 $n=784$)。数据集中捕获的数据包的长度往往是不同的，对于长度大于 784 Byte 的数据包，将对其进行截断。对于长度小于 784 Byte 的数据包，将对其进行拼接，具体拼接做法是将两个数据包进行前后连接，中间插入 50 个 0 Byte 作为间隔。

4)端口号屏蔽与数据归一化。一个会话的数据包具有相同的源 IP 地址、源端口、目的 IP 地址、目的端口。这些相同地址和端口号会大大提高分类的精度，但是在实际应用当中，地址和端口号是随机变化的，因此需要对其进行屏蔽。具体方法是将地址与端口号对应的字节变为 0。最后，将预处理好的流量数据归一化到[0,1]范围，这不仅可以提高准确性，还能提高模型训练的收敛速度。

### 2.3 生成流量图像

经过数据预处理后，将网络流量按照所提出的构图方式生成流量图像。如图 3 所示，不同构图方式下生成的流量图像，每个图像的大小为 784(28×28) Byte。可以明显看到，不同构图方式下生成的流量图像有着完全不一样的图形纹路。每种类型的网络流量都生成了 10 000 张流量图像，其中 9 000 张做训练集，1 000 张做测试集。

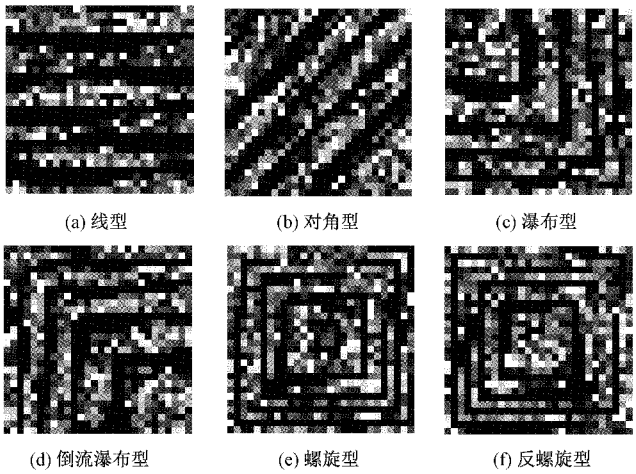


图 3 多种构图方式下的流量图像

### 2.4 基于卷积神经网络的流量图像分类

卷积神经网络，是目前最热门的深度学习算法之一。它在自然语言处理，计算机视觉等领域，都获得了巨大的成功。其在图像处理的应用最为广泛，图像分类，目标识别等[14]。其内部卷积层的功能就是对输入图像的特征进行提取。正一因为卷积神经网络强大的图像特征提取能力，所以选择使用其对流量图像进行分类。

经典的卷积神经网络主要由卷积层、池化层以及全连接层组成。卷积层包含多个卷积核，组成卷积核的每个元素都对应 1 个权重系数和 1 个偏差量，类似于 1 个前馈神经网络的神经元。利用卷积核与输入数据做卷积运算，提取特征。池化层的功能是对卷积层输出的特征图进行特征选择和信息过滤。具体做法就是将特征图中单个点的结果替换为其相邻区域的特征图统计量。其目的是为了减少网络中的参数和计算。均值池化和极大池化是卷积神经网络中最常见的池化方法。全连接层一般位于卷积神经网络的最后，其作用是对卷积层以及池化层提取的特征进行非线性组合以得到输出，它自身不被期望具有特征提取能力。

一般的卷积神经网络都会含有多个全连接层，最后一个全连接层连接输出层。对于图像分类问题，输出层使用逻辑函数或归一化指数函数输出分类标签。

表 2 所示为本次实验所采用的卷积神经网络结构。为了保证结果的可对比性，所有的实验都采用了相同的网络结构与参数。

**表 2　卷积神经网络架构参数**

| 层数 | 类型 | 滤波器/神经元 | 步幅 | 填充 |
|---|---|---|---|---|
| 1 | conv+ReLU | 32 | 1 | same |
| 2 | max pooling | 2 | 2 | same |
| 3 | conv+ReLU | 64 | 1 | same |
| 4 | max pooling | 2 | 2 | same |
| 5 | dense | 1 024 | — | none |
| 6 | dense | 标签的数量 | — | none |
| 7 | softmax | — | — | none |

## 3　实验与结果分析

为了验证多种构图方式的有效性，分别对公开数据集 ISCX VPN-nonVPN 数据集和人工捕获的数据集进行分类。每个数据集分别做了两个对比实验。为了保证实验结果的普遍性，每组实验除了数据集标记方式不同以外，预处理过程与网络模型以参数都保持相同。

### 3.1　评价指标

一共使用 4 个评价指标：准确度(Acc)、精确度(Pre)、召回率(Rec)、$F_1$ 值[15]。每种指标的具体计算方式如式(1)～(4)所示。$TP$ 表示预测为正例而且实际上也是正例；$FP$ 表示预测为正例然而实际上却是负例；$FN$ 表示预测为负例然而实际上却是正例；$TN$ 表示预测为负例而且实际上也是负例。$Acc$ 表示正例和负例中预测正确数量占总数量的比例；$Pre$ 表示预测为正例的样本中预测正确的比例。$Rec$ 表示被预测正确的正例占总实际正例样本的比例。$F_1$ 是对分类结果的综合评价指标。

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Pre = \frac{TP}{TP + FP} \tag{2}$$

$$Rec = \frac{TP}{TP + FN} \tag{3}$$

$$F_1 = \frac{2Pre \times Rec}{Pre + Rec} \tag{4}$$

### 3.2　ISCX 数据集实验结果

按照前文介绍的标记方式，ISCX VPN-nonVPN 数据集分别被标记了 14 类与 22 类。使用完全相同的实验方法对 14 类与 22 类的流量数据进行处理。

表 3 所示为在不同构图模式下对 14 类流量数据进行分类的实验结果。表中记录了每种流量类型分类的的 3 个性能指标，召回率、精确度以及总精确度。表中加黑的数据表示对同一种流量类别 3 个性能指标的最佳值。例如，对于 Chat 类，在对角型的构图方式下，Rec 达到最佳值，在瀑布型的构图方式下，Pre 达到最佳值。从总精确度来看，对角型和瀑布型的结果高于线型，其余 3 种低于线型。其中，对角型的总精度最高，达到 98%，比线型搞了 0.8%，这对于分类精度都高达 95%以上的情况下，也是一个十分可观的提升。从召回率和精确度上来看，最佳值在对角型中出现的频率最高，其次是瀑布型，而在线型中没有出现。这证明了这些特殊的构图模式可以增加数据包字节流之间的相关性，提升分类精度。图 4 是对 22 类流量数据进行分类的实验结果。可以看到在 4 项评价指标上，瀑布型构图方式明显高于线型方式，这表明了在 22 类数据标签下，瀑布型构图方式是更好的选择。

**表 3　ISCX 数据集 14 类分类结果**

| 种类 | 线型 | | 对角型 | | 瀑布型 | | 倒流瀑布型 | | 螺旋型 | | 反螺旋型 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Rec | Pre | Rec | Pre | Rec | Pre | Rec | Pre | Rec | Pre | Rec | Pre |
| Chat | 0.931 | 0.949 | **0.981** | 0.948 | 0.956 | **0.972** | 0.916 | 0.968 | 0.909 | 0.922 | 0.908 | 0.930 |
| Email | 0.893 | 0.940 | 0.885 | **0.987** | 0.923 | 0.959 | **0.929** | 0.904 | 0.817 | 0.869 | 0.888 | 0.940 |
| Filetransfer | 0.979 | 0.966 | 0.991 | **0.984** | **0.992** | 0.983 | 0.988 | 0.980 | 0.890 | 0.967 | 0.971 | 0.951 |
| P2P | 0.996 | 0.986 | 0.998 | **0.996** | 0.998 | 0.993 | 0.992 | 0.982 | **0.999** | 0.984 | 0.991 | 0.971 |
| Streaming | 0.993 | 0.996 | **0.999** | **0.997** | 0.996 | 0.996 | 0.999 | 0.994 | 0.990 | 0.990 | 0.996 | 0.988 |
| VoIP | 0.979 | 0.939 | 0.969 | 0.961 | 0.961 | 0.950 | 0.917 | **0.969** | 0.932 | 0.888 | **0.983** | 0.939 |
| VPN-Chat | 0.980 | 0.956 | **0.991** | 0.969 | 0.973 | **0.973** | 0.968 | 0.951 | 0.966 | 0.934 | 0.968 | 0.967 |
| VPN-Email | 0.960 | 0.977 | 0.949 | **0.986** | 0.964 | 0.972 | 0.959 | 0.985 | **0.979** | 0.956 | 0.968 | 0.975 |
| VPN-Filetransfer | 0.987 | 0.992 | 0.977 | **0.993** | **0.991** | 0.981 | 0.982 | 0.980 | 0.974 | 0.966 | 0.985 | 0.989 |
| VPN-P2P | 0.996 | 0.984 | **0.998** | **0.986** | 0.996 | 0.984 | 0.990 | **0.986** | 0.996 | 0.982 | 0.994 | **0.986** |
| VPN-Streaming | 0.998 | 0.995 | 0.998 | **1.000** | 0.999 | **1.000** | 0.998 | 0.998 | 0.999 | 0.989 | **1.000** | 0.996 |
| VPN-VoIP | 0.911 | 0.963 | **0.944** | 0.955 | 0.926 | 0.941 | 0.940 | 0.887 | 0.781 | 0.899 | 0.913 | **0.970** |
| 总精度 | 0.972 | | **0.981** | | 0.977 | | 0.968 | | 0.951 | | 0.969 | |

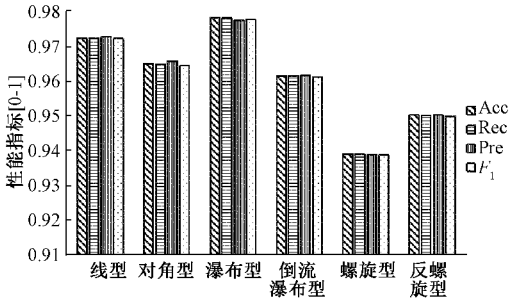


图 4　ISCX 数据集 22 类分类结果

### 3.3　人工捕获数据集实验结果

通过对不同 VPN 下的加密流量进行分类，来证明提出方法的有效性。对于 4 种 VPN 和 8 种 APP，一共有 27 类流量数据。分别做了两个对比实验。实验 1，对 4 种 VPN 下 4 款国外 APP 进行分类。实验 2 在实验 1 基础上，加入了国内 APP 的流量数据。

表 4 所示为对 16 类的国外 APP 的流量数据分类结果。从总精度来看，对角型、瀑布型和倒流瀑布型均高于线型。对角型的总精度最高，达到 93.5%，比线型高了 1%。从每一类流量数据的最佳 Rec 和 Pre 的分布来看，对角型、瀑布型和倒流瀑布型占的数量较多，线型占的数量较少。图 5 是对所有 VPN 与 APP 的 27 类流量数据的分类结果。从整体上来看，由于多加入了 11 类流量数据，每种构图方式的分类结果较 16 类的分类结果都有一点的下降。但下降幅度小于 1%，这是在可接受范围之内的。

**表 4　人工捕获数据集 16 类分类结果**

| 种类 | 线型 | | 对角型 | | 瀑布型 | | 倒流瀑布型 | | 螺旋型 | | 反螺旋型 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| e-facebook | 0.875 | 0.943 | 0.885 | 0.936 | **0.894** | 0.929 | 0.884 | **0.946** | 0.878 | 0.921 | 0.878 | 0.931 |
| e-telegram | 0.983 | 0.769 | 0.956 | **0.851** | **0.975** | 0.813 | 0.970 | 0.812 | 0.960 | 0.800 | 0.984 | 0.777 |
| e-twitter | 0.943 | 0.886 | 0.942 | 0.892 | 0.940 | **0.910** | **0.945** | 0.898 | 0.933 | 0.894 | 0.927 | 0.904 |
| e-youtube | 0.694 | **0.969** | **0.827** | 0.933 | 0.769 | 0.955 | 0.767 | 0.962 | 0.756 | 0.912 | 0.708 | 0.961 |
| p-facebook | 0.996 | 0.988 | **0.999** | 0.980 | 0.998 | **0.994** | 0.998 | 0.992 | 0.977 | 0.966 | 0.990 | 0.981 |
| p-telegram | 0.997 | 0.995 | 0.994 | 0.995 | **1.000** | **0.998** | 0.999 | **0.998** | 0.993 | 0.977 | 0.999 | 0.990 |
| p-twitter | 0.987 | 0.985 | **0.992** | 0.986 | 0.985 | 0.979 | 0.989 | **0.990** | 0.951 | 0.963 | 0.988 | 0.982 |
| p-youtube | 0.980 | 0.970 | 0.970 | 0.982 | **0.995** | 0.956 | 0.989 | **0.983** | 0.953 | 0.969 | 0.954 | 0.954 |
| s-facebook | 0.887 | **0.889** | **0.943** | 0.854 | 0.906 | 0.831 | 0.925 | 0.857 | 0.795 | 0.797 | **0.943** | 0.764 |
| s-telegram | 0.927 | 0.917 | 0.889 | **0.936** | 0.919 | 0.914 | **0.932** | 0.928 | 0.860 | 0.866 | 0.867 | 0.926 |
| s-twitter | **0.949** | 0.908 | 0.937 | 0.951 | 0.882 | 0.924 | 0.880 | **0.952** | 0.802 | 0.784 | 0.847 | 0.942 |
| s-youtube | **0.883** | 0.936 | 0.872 | 0.946 | 0.870 | **0.954** | 0.869 | 0.944 | 0.805 | 0.930 | 0.863 | 0.913 |
| v-facebook | 0.880 | 0.908 | 0.896 | **0.920** | 0.879 | 0.912 | 0.917 | 0.904 | 0.898 | 0.859 | **0.941** | **0.895** |
| v-telegram | 0.953 | 0.979 | 0.967 | **0.982** | 0.923 | 0.979 | **0.970** | 0.981 | 0.937 | 0.947 | 0.952 | 0.977 |
| v-twitter | 0.967 | 0.954 | 0.962 | **0.955** | **0.972** | 0.946 | 0.968 | 0.918 | 0.949 | 0.886 | 0.967 | 0.855 |
| v-youtube | 0.904 | 0.870 | 0.923 | 0.881 | **0.991** | 0.868 | 0.903 | 0.881 | 0.847 | 0.853 | 0.877 | **0.922** |
| 总精度 | 0.925 | | **0.935** | | 0.926 | | 0.932 | | 0.893 | | 0.912 | |

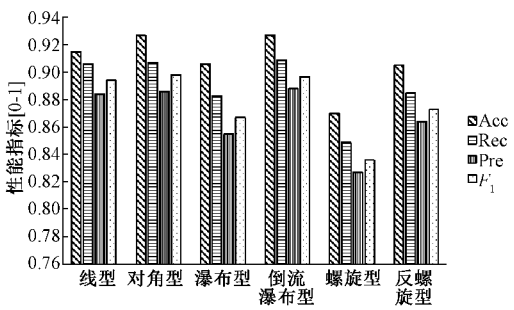


图 5　人工捕获数据集 27 类分类结果

## 4　结　　论

本文提出了一种基于多种构图方式的加密流量分类方法，实现对多种 VPN 的加密流量进行分类。构造了一个基于多种 VPN 的热门 APP 加密流量数据集，并使用了多种构图方式（对角型、瀑布型、螺旋型等）将加密流量数据包转换成流量图像，输入到卷积神经网络进行分类。最终的结果表明，使用对角型或瀑布型构图方式的分类精度往往会比传统的线型构图方式的分类精度要高。这种特殊的构图方式可以加强流量图像中像素点的相关性，增加流量图像中的图像特征，最终提升加密流量的分类精度。

## 参考文献

[1]　杨晓敏. 改进灰狼算法优化支持向量机的网络流量预测[J]. 电子测量与仪器学报，2021，35(3)：211-217.

[2]　WANG W. HAST-IDS：Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection[J]. IEEE Access，2018，

6: 1792-1806.

[3] WANG Q, YAHYAVI A, KEMME B, et al. I know what you did on your smartphone: Inferring app usage over encrypted data traffic[C]. 2015 IEEE Conference on Communications and Network Security, 2015: 433-441.

[4] HE Y, LI W. Image-based encrypted traffic classification with convolution neural networks[C]. 2020 IEEE Fifth International Conference on Data Science in Cyberspace, 2020:271-278.

[5] WANG W, ZHU M, WANG J, et al. End-to-end encrypted traffic classification with one-dimensional convolution neural networks [C]. 2017 IEEE International Conference on Intelligence and Security Informatics, 2017:43-48.

[6] LOTFOLLAHI M, JAFARI S M, ZADE R S H, et al. Deep packet: A novel approach for encrypted traffic classification using deep learning[J]. Soft Comput, 2020, 24(3):1999-2012.

[7] SHAPIRA T, SHAVITT Y. FlowPic: Encrypted internet traffic classification is as easy as image recognition [C]. IEEE Conference on Computer Communications Workshops, 2019:680-687.

[8] ACETO G, CIUONZO D, MONTIERI A, et al. Mobile encrypted traffic classification using deep learning: Experimental evaluation, lessons learned, and challenges[J]. IEEE Transactions on Network and Service Management, 2019, 16(2):445-458.

[9] LOPEZ-MARTIN M, CARRO B, SANCHEZ A, et al. Network traffic classifier with convolutional and recurrent neural networks for internet of things[J]. IEEE Access, 2017, 5:18042-18050.

[10] SALEH I, JI H. Network traffic images: A deep learning approach to the challenge of internet traffic classification[C]. Annual Computing and Communication Workshop and Conference, 2020:0329-0334.

[11] MULIUKHA V A, LABOSHIN L U. Analysis and classification of encrypted network traffic using machine learning[C]. International Conference on Soft Computing and Measurements, 2020:194-197.

[12] HABIBI L A, DRAPER-GIL G, MAMUN M, et al. Characterization of encrypted and VPN traffic using time-related features [C]. International Conference on Information Systems Security and Privacy, 2016:407-414.

[13] WANG P, LI S, YE F, et al. PacketCGAN: Exploratory study of class imbalance for encrypted traffic classification using CGAN [C]. IEEE International Conference on Communications, 2020: 1-7.

[14] KRIZHEVSKY A, SUTSKEVER I. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6):84-90.

[15] ILIYASU A S, DENG H. Semi-supervised encrypted traffic classification with deep convolutional generative adversarial networks [J]. IEEE Access, 2020, 8: 118-126.

## 作者简介

**朱文斌**，硕士，主要研究方向为加密流量分类、大数据。
E-mail: zhuwenbin@shu.edu.cn

**马秀丽**，博士，副教授，主要研究方向为模式识别、人工智能。
E-mail: xlma@shu.edu.cn