(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 120415755 A
(43)申请公布日 2025.08.01

(21)申请号 202410135624.8

(22)申请日 2024.01.31

(71)申请人 中国科学院计算机网络信息中心
地址 100190 北京市海淀区中关村南四街4号院内2号楼

(72)发明人 宫良一 龙春 杨帆 魏金侠 付豪

(74)专利代理机构 北京知舟专利事务所(普通合伙) 11550
专利代理师 郭辗

(51)Int.Cl.
H04L 9/40(2022.01)

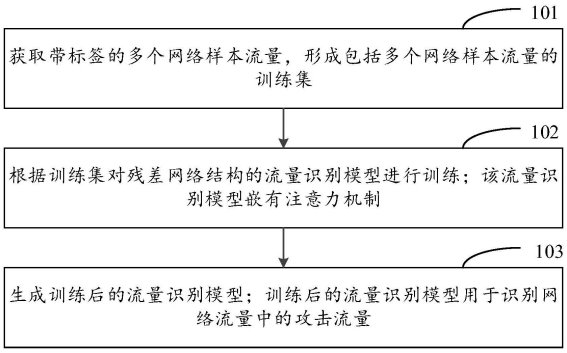
权利要求书2页 说明书14页 附图7页

(54)发明名称

攻击流量的识别方法、装置、电子设备及存储介质

(57)摘要

本发明提供了一种攻击流量的识别方法、装置、电子设备及存储介质,其中,该方法包括:获取带标签的多个网络样本流量,形成包括多个所述网络样本流量的训练集;根据所述训练集对残差网络结构的流量识别模型进行训练;所述流量识别模型嵌有注意力机制;生成训练后的流量识别模型;所述训练后的流量识别模型用于识别网络流量中的攻击流量。通过本发明实施例提供的方案,可以更好地学习到输入数据和输出数据之间的差异,能够提升流量识别模型的性能和训练效果;并且,基于注意力机制来增强流量识别模型对输入数据中重要特征的关注,并抑制不重要的特征,从而可以提高流量识别模型的表示能力和性能,能够更加准确地识别出新型的攻击流量。



CN 120415755 A

1.一种攻击流量的识别方法,其特征在于,包括:

获取带标签的多个网络样本流量,形成包括多个所述网络样本流量的训练集;

根据所述训练集对残差网络结构的流量识别模型进行训练;所述流量识别模型嵌有注意力机制;

生成训练后的流量识别模型;所述训练后的流量识别模型用于识别网络流量中的攻击流量。

2.根据权利要求1所述的方法,其特征在于,所述流量识别模型包括依次串接的至少一个残差部;

所述残差部包括依次串接的一个卷积块以及至少一个恒等块;所述卷积块嵌有注意力机制。

3.根据权利要求2所述的方法,其特征在于,所述卷积块包括:位于主路依次串接的至少一个注意力残差单元,位于支路的支路卷积层,以及第一相加层;所述注意力残差单元包括:第一卷积层、第一激活层、第一注意力层和相乘层;

所述支路卷积层用于,对所述卷积块的输入数据进行卷积处理,生成不同通道数的输出数据;

所述第一卷积层用于,对所述注意力残差单元的输入数据进行卷积处理;

所述第一激活层用于,基于激活函数对所述第一卷积层的输出数据进行非线性处理;

所述第一注意力层用于,确定所述第一激活层的输出数据中每个特征的注意力权重;

所述相乘层用于,对所述第一注意力层输出的注意力权重与所述第一激活层的输出数据进行相乘处理;

所述第一相加层用于,对所述主路的输出数据与所述支路的输出数据进行相加处理。

4.根据权利要求3所述的方法,其特征在于,所述恒等块包括:位于主路依次串接的至少一个恒等残差单元,以及第二相加层;所述恒等残差单元包括:第二卷积层和第二激活层;

所述第二卷积层用于,对所述恒等残差单元的输入数据进行卷积处理;

所述第二激活层用于,基于激活函数对所述第二卷积层的输出数据进行非线性处理;

所述第二相加层用于,对所述主路的输出数据与所述恒等块的输入数据进行相加处理。

5.根据权利要求2所述的方法,其特征在于,所述流量识别模型还包括注意力部;

所述注意力部与最后一个残差部相连,用于对所述最后一个残差部的输出数据引入注意力机制,生成上下文向量,并连接所述上下文向量以及所述最后一个残差部的输出数据,生成所述流量识别模型的输出数据。

6.根据权利要求5所述的方法,其特征在于,所述注意力部包括:平均池化层、重构层、第二注意力层、自定义层、第一点积层、第三激活层、第二点积层、展平层、第一全连接层、数据链接层和第二全连接层;

所述平均池化层用于,对所述最后一个残差部的输出数据进行平均池化处理;

所述重构层用于,对所述平均池化层的输出数据进行重构变形;

所述第二注意力层用于,确定所述重构层的输出数据中每个特征的注意力权重;

所述自定义层用于,提取所述重构层的输出数据中最后一个时间步的隐藏状态;

所述第一点积层用于,对所述第二注意力层输出的注意力权重与所述自定义层输出的隐藏状态进行点积处理,生成注意力分数;

所述第三激活层用于,基于激活函数对所述第一点积层输出的注意力分数进行非线性处理;

所述第二点积层用于,对所述第三激活层的输出数据与所述自定义层输出的隐藏状态进行点积处理,生成上下文向量;

所述展平层用于,对所述平均池化层的输出数据展平为一维向量;

所述第一全连接层用于,对所述展平层的输出数据进行特征提取;

所述数据链接层用于,连接所述第二点积层输出的上下文向量以及所述第一全连接层的输出数据;

所述第二全连接层用于,对所述数据链接层的输出数据进行特征提取,输出所述流量识别模型的输出数据。

7.根据权利要求1至6中任一项所述的方法,其特征在于,所述根据所述训练集对残差网络结构的流量识别模型进行训练,包括:

在测试阶段确定预测标签的F1分数;

根据所述F1分数对所述流量识别模型的权重进行更新,且更新后的权重满足:

$$w_{new} = w_{old} + \frac{1}{F_1 + a};$$

其中,w_{new}表示更新后的权重,w_{old}表示更新前的原始权重,F_1表示所述F1分数,a为防止除零的极小值。

8.一种攻击流量的识别装置,其特征在于,包括:

获取模块,用于获取带标签的多个网络样本流量,形成包括多个所述网络样本流量的训练集;

训练模块,用于根据所述训练集对残差网络结构的流量识别模型进行训练;所述流量识别模型嵌有注意力机制;

处理模块,用于生成训练后的流量识别模型;所述训练后的流量识别模型用于识别网络流量中的攻击流量。

9.一种电子设备,包括总线、收发器、存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述收发器、所述存储器和所述处理器通过所述总线相连,其特征在于,所述计算机程序被所述处理器执行时实现如权利要求1至7中任一项所述的攻击流量的识别方法中的步骤。

10.一种计算机可读存储介质,其上存储有计算机程序,其特征在于,所述计算机程序被处理器执行时实现如权利要求1至7中任一项所述的攻击流量的识别方法中的步骤。

攻击流量的识别方法、装置、电子设备及存储介质

技术领域

[0001] 本发明涉及网络安全技术领域,具体而言,涉及一种攻击流量的识别方法、装置、电子设备及存储介质。

背景技术

[0002] 随着当前网络通信技术的发展,在为人们日常生活提供便利的同时,所积累的大量包含用户隐私的敏感数据也成为了攻击者的入侵目标。当前,对互联网的攻击常分为五大类:DoS/DDoS攻击、信息收集、中间人攻击、注入攻击以及恶意软件攻击。例如,许多基于互联网、物联网的组织,包括医疗保健、政府和研究机构,在研究与新型网络相关的威胁和攻击时,都面临着潜在的障碍。

[0003] 传统基于知识的入侵检测系统(IDS)在面对复杂网络和新型攻击时难以适应,因为这些新颖的攻击和内部人员的恶意行为可能具有的可疑痕迹较少。因此,现代的入侵检测系统需要采用智能的数据驱动方法,利用机器学习和数据分析技术,从大量的数据中学习模式和统计规律,并识别异常情况,从而检测入侵活动,以能够检测可能感染新型网络的恶意攻击。

[0004] 面对复杂网络和新型攻击时难以适应新颖的攻击和内部人员的恶意行为

[0005] 现代攻击流量入侵检测方法近似的实现方案有:(1)基于卷积神经网络结构,包括编码器网络,解码器网络和像素分类层,可以将低分辨率编码器特征映射转换为全输入分辨率特征映射以进行像素分类。(2)在安全的自动两级入侵检测系统(SATIDS)的基础上改进了长短期记忆(LSTM)网络,该系统能够区分攻击和正常的网络流量,并且能够识别不同类型的攻击,并对其进行细分。(3)基于联邦混合模型驱动的IDS框架的物联网(IoT)和工业物联网(IIoT),称为F-BIDS。对Edge-IIoTset数据集15分类得出的结果为90.91%,6分类为90.93%。

[0006] 目前基于深度学习的入侵检测方法虽然可以一定程度上适应新型攻击,但在灵活性和适用性上存在着一定缺陷,从正常流量中捕捉异常行为时误报率较高。

发明内容

[0007] 为解决现有存在的技术问题,本发明实施例提供一种攻击流量的识别方法、装置、电子设备及存储介质。

[0008] 第一方面,本发明实施例提供了一种攻击流量的识别方法,包括:

[0009] 获取带标签的多个网络样本流量,形成包括多个所述网络样本流量的训练集;

[0010] 根据所述训练集对残差网络结构的流量识别模型进行训练;所述流量识别模型嵌有注意力机制;

[0011] 生成训练后的流量识别模型;所述训练后的流量识别模型用于识别网络流量中的攻击流量。

[0012] 在一些可选的实施方式中,所述流量识别模型包括依次串接的至少一个残差部;

[0013] 所述残差部包括依次串接的一个卷积块以及至少一个恒等块;所述卷积块嵌有注意力机制。

[0014] 在一些可选的实施方式中,所述卷积块包括:位于主路依次串接的至少一个注意力残差单元,位于支路的支路卷积层,以及第一相加层;所述注意力残差单元包括:第一卷积层、第一激活层、第一注意力层和相乘层;

[0015] 所述支路卷积层用于,对所述卷积块的输入数据进行卷积处理,生成不同通道数的输出数据;

[0016] 所述第一卷积层用于,对所述注意力残差单元的输入数据进行卷积处理;

[0017] 所述第一激活层用于,基于激活函数对所述第一卷积层的输出数据进行非线性处理;

[0018] 所述第一注意力层用于,确定所述第一激活层的输出数据中每个特征的注意力权重;

[0019] 所述相乘层用于,对所述第一注意力层输出的注意力权重与所述第一激活层的输出数据进行相乘处理;

[0020] 所述第一相加层用于,对所述主路的输出数据与所述支路的输出数据进行相加处理。

[0021] 在一些可选的实施方式中,所述恒等块包括:位于主路依次串接的至少一个恒等残差单元,以及第二相加层;所述恒等残差单元包括:第二卷积层和第二激活层;

[0022] 所述第二卷积层用于,对所述恒等残差单元的输入数据进行卷积处理;

[0023] 所述第二激活层用于,基于激活函数对所述第二卷积层的输出数据进行非线性处理;

[0024] 所述第二相加层用于,对所述主路的输出数据与所述恒等块的输入数据进行相加处理。

[0025] 在一些可选的实施方式中,所述流量识别模型还包括注意力部;

[0026] 所述注意力部与最后一个残差部相连,用于对所述最后一个残差部的输出数据引入注意力机制,生成上下文向量,并连接所述上下文向量以及所述最后一个残差部的输出数据,生成所述流量识别模型的输出数据。

[0027] 在一些可选的实施方式中,所述注意力部包括:平均池化层、重构层、第二注意力层、自定义层、第一点积层、第三激活层、第二点积层、展平层、第一全连接层、数据链接层和第二全连接层;

[0028] 所述平均池化层用于,对所述最后一个残差部的输出数据进行平均池化处理;

[0029] 所述重构层用于,对所述平均池化层的输出数据进行重构变形;

[0030] 所述第二注意力层用于,确定所述重构层的输出数据中每个特征的注意力权重;

[0031] 所述自定义层用于,提取所述重构层的输出数据中最后一个时间步的隐藏状态;

[0032] 所述第一点积层用于,对所述第二注意力层输出的注意力权重与所述自定义层输出的隐藏状态进行点积处理,生成注意力分数;

[0033] 所述第三激活层用于,基于激活函数对所述第一点积层输出的注意力分数进行非线性处理;

[0034] 所述第二点积层用于,对所述第三激活层的输出数据与所述自定义层输出的隐藏

状态进行点积处理,生成上下文向量;

[0035] 所述展平层用于,对所述平均池化层的输出数据展平为一维向量;

[0036] 所述第一全连接层用于,对所述展平层的输出数据进行特征提取;

[0037] 所述数据链接层用于,连接所述第二点积层输出的上下文向量以及所述第一全连接层的输出数据;

[0038] 所述第二全连接层用于,对所述数据链接层的输出数据进行特征提取,输出所述流量识别模型的输出数据。

[0039] 在一些可选的实施方式中,所述根据所述训练集对残差网络结构的流量识别模型进行训练,包括:

[0040] 在测试阶段确定预测标签的F1分数;

[0041] 根据所述F1分数对所述流量识别模型的权重进行更新,且更新后的权重满足:

[0042] $$w_{new} = w_{old} + \frac{1}{F_1 + a};$$

[0043] 其中,w_{new}表示更新后的权重,w_{old}表示更新前的原始权重,F_1表示所述F1分数,a为防止除零的极小值。

[0044] 第二方面,本发明实施例还提供了一种攻击流量的识别装置,包括:

[0045] 获取模块,用于获取带标签的多个网络样本流量,形成包括多个所述网络样本流量的训练集;

[0046] 训练模块,用于根据所述训练集对残差网络结构的流量识别模型进行训练;所述流量识别模型嵌有注意力机制;

[0047] 处理模块,用于生成训练后的流量识别模型;所述训练后的流量识别模型用于识别网络流量中的攻击流量。

[0048] 第三方面,本发明实施例提供了一种电子设备,包括总线、收发器、存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述收发器、所述存储器和所述处理器通过所述总线相连,所述计算机程序被所述处理器执行时实现上述任意一项所述的攻击流量的识别方法中的步骤。

[0049] 第四方面,本发明实施例还提供了一种计算机可读存储介质,其上存储有计算机程序,所述计算机程序被处理器执行时实现上述任意一项所述的攻击流量的识别方法中的步骤。

[0050] 本发明实施例提供的攻击流量的识别方法、装置、电子设备及存储介质,采用残差网络结构的流量识别模型,利用残差网络中的跳连接,可以更好地学习到输入数据和输出数据之间的差异,能够提升流量识别模型的性能和训练效果;并且,基于注意力机制来增强流量识别模型对输入数据中重要特征的关注,并抑制不重要的特征,从而可以提高流量识别模型的表示能力和性能,能够更加准确地识别出新型的攻击流量。在面对复杂网络和新型攻击时,可以适应新型攻击或内部人员的恶意行为,对攻击流量的识别效果较好,分类精度较高。

[0051] 通过在卷积块中嵌入注意力机制,使模型能够动态地关注输入中得重要特征,并通过残差连接保持信息的传递。这种注意力机制的引入有助于提高模型的表达能力和性能,使其能够更好地适应不同的数据分布和任务要求。并且,通过引入注意力机制,通过一

系列操作和权重计算,生成上下文向量,将上下文向量和之前的模型输出进行连接,以融合注意力信息和其他特征,使模型能够更智能地处理输入数据,并提高模型在各种任务上的性能。这种注意力机制允许模型有选择性地关注输入数据中的不同部分,从而更好地理解和利用输入数据的信息。通过调整训练中预测标签类别的权重,可以保证预测标签在少量样本的情况下,也能预测良好,能够优化模型的性能。

附图说明

[0052] 为了更清楚地说明本发明实施例或背景技术中的技术方案,下面将对本发明实施例或背景技术中所需要使用的附图进行说明。

[0053] 图1示出了本发明实施例所提供的一种攻击流量的识别方法的流程图;

[0054] 图2示出了本发明实施例所提供的流量识别模型的一种结构示意图;

[0055] 图3示出了本发明实施例所提供的流量识别模型中,卷积块的一种结构示意图;

[0056] 图4示出了本发明实施例所提供的流量识别模型中,恒等块的一种结构示意图;

[0057] 图5示出了本发明实施例所提供的流量识别模型的另一种结构示意图;

[0058] 图6示出了本发明实施例所提供的流量识别模型中,注意力部的一种结构示意图;

[0059] 图7示出了本发明实施例所提供的一种攻击流量的识别装置的结构示意图;

[0060] 图8示出了本发明实施例所提供的一种用于执行攻击流量的识别方法的电子设备的结构示意图。

具体实施方式

[0061] 下面结合本发明实施例中的附图对本发明实施例进行描述。

[0062] 图1示出了本发明实施例所提供的一种攻击流量的识别方法的流程图。如图1所示,该方法包括步骤101至步骤103。

[0063] 步骤101:获取带标签的多个网络样本流量,形成包括多个网络样本流量的训练集。

[0064] 本实施例中,为实现模型训练,需要设置训练集。具体地,可以获取多个具有标签的网络流量,为方便描述,将具有标签的网络流量称为网络样本流量;其中,该标签具体表示网络样本流量是否为攻击流量,即网络样本流量的标签分为攻击流量和正常流量。在确定多个网络样本流量,即可形成相应的训练集。

[0065] 其中,由于在训练过程中还需要对模型进行测试,故也需要设置相应的测试集。本实施例中,可以将获取的多个网络样本流量按照合适的比例(例如8:2),分为训练集和测试集。

[0066] 可选地,为保证网络样本流量能够适合流量识别模型的数据格式,一般需要对获取到的网络样本流量进行预处理,最终形成训练集。

[0067] 具体地,在获取到包含多个网络样本流量的数据集后,可以首先使用标签编码方法将网络样本流量中非数值型的分类变量值转换为数值。下一阶段,进行数据清理,从数据集中删除丢失和无意义以及重复的数据,以提高数据的质量和模型的性能。最后,对数据进行清理后,数据集可以被分为训练集和测试集。

[0068] 在数据清洗过程中,可以使用类别权重调整技术(例如,最小-最大缩放技术)重新

调整所有列(每一列对应网络样本流量的同一类数据),实现数据归一化,保证它们都在相同比例上,例如将数据归一化至[0,1]。该技术的数学模型定义如下:

[0069] $$x' = \frac{x - \min(X)}{\max(X) - \min(X)} \quad (1)$$

[0070] 其中,x表示数据集中某类数据X中的一个样本,x'表示对样本x调整后所确定的值;min(X)表示该类数据X中所有样本的最小值,max(X)表示该类数据X中所有样本的最大值。使用类别权重调整技术对数据集进行调整,可以提高模型对样本数量较少类别的关注度,平衡各个类别。

[0071] 步骤102:根据训练集对残差网络结构的流量识别模型进行训练;该流量识别模型嵌有注意力机制。

[0072] 本实施例中,在确定训练集之后,即可基于该训练集对流量识别模型进行训练,以训练得到能够进行流量识别的流量识别模型。该流量识别模型可以是深度强化自学习模型。具体地,本实施例基于残差网络结构搭建该流量识别模型,即该流量识别模型主体框架为残差网络(ResNet)。

[0073] 用于识别攻击流量的传统模型,其一般也存在对新型攻击误报率较高的问题;为能够提高识别率,达到更高的分类精度,一般会增加模型层数,例如增加卷积层的数量;但这种改进只在一定程度上有效,还容易导致梯度消失或爆炸的问题,从而限制了模型层数,导致模型分类精度不可能很高。而在本实施例中,采用残差网络结构的流量识别模型,可以有效解决梯度消失或爆炸的问题,提高模型的分类性能。

[0074] 具体地,残差网络采用跳连接的方法,通过跳过一些层,创建残差块,将前面层的激活连接到后续层。通过残差连接,网络可以更好地学习到输入数据和输出数据之间的差异,解决了由于梯度消失或爆炸,导致模型性能下降等问题,从而可以提升流量识别模型的性能和训练效果。

[0075] 并且,在该流量识别模型中,还引入了注意力机制,从而可以基于注意力机制来增强流量识别模型对输入数据中重要特征的关注,并抑制不重要的特征,从而可以提高流量识别模型的表示能力和性能,使得流量识别模型能够比较好地提取出新型攻击流量中的特点,从而可以更加准确地识别出新型的攻击流量。

[0076] 步骤103:生成训练后的流量识别模型;训练后的流量识别模型用于识别网络流量中的攻击流量。

[0077] 本实施例中,利用训练集对上述残差网络结构的流量识别模型进行训练,即可得到训练后的流量识别模型,进而基于该训练后的流量识别模型进行攻击流量识别;该流量识别模型相当于基于强化学习的入侵检测系统,可以有效解决互联网攻击流量的识别问题。具体地,在上述步骤103之后,该方法还可以包括:将待识别的网络流量输入至该训练后的流量识别模型,以确定该网络流量是否为攻击流量。

[0078] 本实施例提供的攻击流量的识别方法,采用残差网络结构的流量识别模型,利用残差网络中的跳连接,可以更好地学习到输入数据和输出数据之间的差异,能够提升流量识别模型的性能和训练效果;并且,基于注意力机制来增强流量识别模型对输入数据中重要特征的关注,并抑制不重要的特征,从而可以提高流量识别模型的表示能力和性能,能够

更加准确地识别出新型的攻击流量。在面对复杂网络和新型攻击时,可以适应新型攻击或内部人员的恶意行为,对攻击流量的识别效果较好,分类精度较高。

[0079] 在一些可选的实施方式中,图2示出了流量识别模型的一种结构示意图,如图2所示,该流量识别模型包括依次串接的至少一个残差部210。本实施例中,将流量识别模型中实现残差连接的结构称为残差部210,该残差部210的数量可以为一个,也可以为多个;一般情况下,为提高识别精度,会采用多个残差部210,图2以流量识别模型包括一个残差部210为例示出。

[0080] 如图2所示,该流量识别模型还包括初始的卷积处理结构,该卷积处理结构包括零填充层201、卷积层202、激活层203和最大池化层204,基于该卷积处理结构对流量识别模型的输入数据(例如,训练过程中的网络样本流量,或者识别过程中待识别的网路流量等)进行初步卷积处理,并向残差部210提供处理后的数据。

[0081] 具体地,该零填充层(Zero Padding)201用于对输入数据进行零填充,以增加输入数据的边界,这是为了确保在卷积操作中能够捕获输入数据的边缘信息。例如,可以对输入数据的周围添加一圈零像素,实现零填充。其中,该零填充层201可以是一维(1D)的结构。

[0082] 卷积层(Conv)202用于对零填充后的数据进行卷积处理。例如,该卷积层202可以使用大小为7的卷积核,即卷积核大小为7×7,其数量可以为64。本实施例中,该卷积层202也可以是一维(1D)的结构。例如,该卷积层202的卷积操作可以将输入数据的通道数从3减少到1,以便后续的注意力权重计算。卷积层是一种在深度学习中广泛应用的层类型,它在提取基本特征并提高分类准确行方面具有可靠性;卷积层与所提出的模型结合使用,能够从数据中提取空间表示。

[0083] 激活层(Activation)203用于基于激活函数对卷积层202的处理结果进行非线性处理;其中,该激活函数例如可以为ReLU函数等,本实施例对此不做限定。

[0084] 最大池化层(Max Pooling)204用于对激活层203输出的数据进行最大池化处理,提取出其中的特征,并向残差部210输出相应的数据。其中,该最大池化层204也可以是一维(1D)的结构。

[0085] 本实施例中,如图2所示,该残差部210包括依次串接的一个卷积块300以及至少一个恒等块400;即卷积块300可以向恒等块400输出相应的数据,并且,恒等块400的数量可以为多个,且多个恒等块400依次串联。如图2所示,该残差部210包括依次串联的一个个卷积块300以及两个恒等块400。

[0086] 其中,卷积块300和恒等块400都是一种残差块。这些块用于构建深度残差网络,其中卷积块300用于学习新的特征,而恒等块400用于保留原始特征并传递到下一层。例如,该流量识别模型可以采用ResNet50的网络结构,其具体可以包括四个残差部210,每个残差部210均包括一个卷积块300和多个恒等块400。

[0087] 具体地,卷积块300是输入数据与输出数据之间通道数不同的残差块。卷积块300通常包括一系列卷积层和激活函数等组成,用于提取输入数据中的特征,通过多个卷积层的组合来捕获不同层次的特征信息。其中,卷积块300可以包括不同尺寸的卷积核,以便捕获局部和全局特征。

[0088] 恒等块400是输入数据与输出数据之间通道数相同的残差块,其是一种特殊的残差块,用于执行恒等映射,恒等映射表示输入数据与输出数据之间没有显著的变化或特征

提取。恒等块400通过跳跃连接,可以将输入特征映射直接添加到输出特征映射。

[0089] 本实施例中,在卷积块的设计中,为卷积块300引入注意力机制,从而提高卷积块学习新特征的能力,可以提高模型的表示能力和性能。恒等块400可以采用传统残差网络中的结构。

[0090] 可选地,与传统卷积块结构相似,本实施例引入注意力机制的卷积块300也包括主路和支路。参见图3所示,该卷积块300包括:位于主路依次串接的至少一个注意力残差单元301,位于支路的支路卷积层302,以及第一相加层303。如图3所示,卷积块300的左侧为主路,右侧为支路;左侧包括多个依次串接的注意力残差单元301,图3以包括两个注意力残差单元301为例示出;右侧支路设有一个卷积层,即支路卷积层302;主路和支路的输出经过相加层后得到该卷积块300最终的输出数据。为方便描述,将该相加层称为第一相加层303。

[0091] 其中,支路卷积层302用于,对卷积块300的输入数据进行卷积处理,生成不同通道数的输出数据,即支路卷积层302通过设置合适数量的卷积核,可以改变卷积块300的输入数据的通道数。可以理解,该卷积块300的输入数据可以是最大池化层204的输出数据,也可以是其他残差部210的输出数据。

[0092] 卷积块300的主路用于基于至少一个注意力残差单元301对卷积块300的输入数据进行特征提取,提取出引入注意力的特征。

[0093] 第一相加层303用于,对主路的输出数据与支路的输出数据进行相加处理,即对主路的输出数据与支路卷积层302的输出数据进行相加处理,从而得到该卷积块300的输出数据。

[0094] 并且,如图3所示,该注意力残差单元301包括:第一卷积层3011、第一激活层3012、第一注意力层3013和相乘层3014。

[0095] 其中,第一卷积层3011用于,对注意力残差单元301的输入数据进行卷积处理;第一激活层3012用于,基于激活函数对第一卷积层3011的输出数据进行非线性处理;第一注意力层3013用于,确定第一激活层3012的输出数据中每个特征的注意力权重;相乘层3014用于,对第一注意力层3013输出的注意力权重与第一激活层3012的输出数据进行相乘处理。

[0096] 具体地,如图3所示,对于卷积块300的输入数据,其可作为第一个注意力残差单元301的输入数据,首先由第一卷积层3011对其进行卷积处理,其可以对输入数据进行降维,这样做的目的是将输入数据的维度与后续注意力权重的维度相匹配,以便进行有效的操作;之后再由第一激活层3012对第一卷积层3011的输出数据进行非线性处理,即对降维后的特征进行处理,从而可以增强其表征能力和非线性特性。其中,在该第一卷积层3011之后可以应用批量归一化层(Batch Normalization,简称BN层),以将输入数据进行规范化,使得均值接近0,方差接近1,从而减少了训练过程中的梯度消失和梯度爆炸问题,有助于模型更稳定地学习特征表示,提高模型的训练速度和性能。由于批量归一化层通常在卷积块或恒等块内部使用,在图3中未示批量归一化层,批量归一化层具体位于卷积操作之后、激活函数之前,即位于第一卷积层3011与第一激活层3012之间。

[0097] 对于该第一激活层3012的输出数据,将其同时输入至第一注意力层3013以及相乘层3014,基于第一注意力层3013对该第一激活层3012的输出数据中每个特征计算注意力权重,最后利用相乘层3014对第一激活层3012的输出数据增加该注意力权重,从而得到加权

的注意力特征,将其作为整个注意力残差单元301的输出数据。

[0098] 例如,该第一注意力层3013具体可以是一个全连接层(Dense),基于全连接层来生成注意力权重。该全连接层可以设有Sigmoid激活函数,经过Sigmoid激活函数的处理,将注意力权重限制在0到1之间。这样,生成的注意力权重可以用来衡量每个输入特征的重要性。之后,基于相乘层3014将这些注意力权重与降维后的特征逐元素相乘,从而可以突出重要的特征,并抑制不重要的特征。其中,Dense层是Keras中的全连接层,本实施例结合适当的激活函数和归一化操作,使得Dense层可以用来生成注意力权重。

[0099] 注意力残差单元301的输出数据可以输入至下一个注意力残差单元301;若当前的注意力残差单元301为最后一个,则其输出数据即可作为整个主路的输出数据,并输入至第一相加层303。

[0100] 本实施例中,卷积块300一般包括多个注意力残差单元301。位于之后的注意力残差单元301(不是第一个的注意力残差单元301),其第一卷积层3011在保留注意力加权特征的同时,可以进一步提取和组合特征表示,即其可以作为是特征的进一步加工和筛选层,有助于模型更好地理解输入数据的复杂结构。

[0101] 其中,在最后一个注意力残差单元301与第一相加层303之间,还可以设有一个卷积层,即最后一个注意力残差单元301的输出数据需要经过该卷积层的卷积处理,再输入至第一相加层303,与支路卷积层302的输出数据进行相加。并且,在第一相加层303之后还可以设置激活层,该激活层与第一激活层3012功能相似;例如,可以通过ReLU激活函数对第一相加层303的输出数据进行非线性处理,以获得卷积块300最终的输出。

[0102] 本实施例中,通过在卷积块300中嵌入注意力机制,使模型能够动态地关注输入中得重要特征,并通过残差连接保持信息的传递。这种注意力机制的引入有助于提高模型的表达能力和性能,使其能够更好地适应不同的数据分布和任务要求。

[0103] 可选地,参见图4所示,该恒等块400包括:位于主路依次串接的至少一个恒等残差单元401,以及第二相加层402;恒等残差单元401包括:第二卷积层4011和第二激活层4012。

[0104] 其中,第二卷积层4011用于,对恒等残差单元401的输入数据进行卷积处理;第二激活层4012用于,基于激活函数对第二卷积层4011的输出数据进行非线性处理;第二相加层402用于,对主路的输出数据与恒等块400的输入数据进行相加处理。

[0105] 本实施例中,第二卷积层4011、第二激活层4012、第二相加层402,与卷积块300的第一卷积层3011、第一激活层3012、第一相加层303工作原理相同,此处不做赘述。

[0106] 其中,恒等块400与卷积块300之间的区别主要在于:恒等块400未引入注意力机制,且其支路不需要设置额外的卷积层,即直接将恒等块400的输入数据作为第二相加层402的一个输入,第二相加层402对主路的输出数据以及该恒等块400的输入数据进行相加处理。

[0107] 并且,与卷积块300相似,在最后一个恒等残差单元401与第二相加层402之间,还可以设有一个卷积层(具体可参见图4所示),即最后一个恒等残差单元401的输出数据需要经过该卷积层的卷积处理,再输入至第二相加层402,与恒等块400的输入数据进行相加。并且,在第二相加层402之后还可以设置激活层(具体可参见图4所示),该激活层与第二激活层4012功能相似;例如,可以通过ReLU激活函数对第二相加层402的输出数据进行非线性处理,以获得恒等块400最终的输出。

[0108] 本实施例中,通过残差连接将主路的输出与输入进行加和操作,这种残差连接可以保持信息的传递和梯度的流动,有助于减轻梯度消失问题,并能够促进模型的训练和优化。

[0109] 在一些可选的实施方式中,该流量识别模型还可以引入注意力函数,以增强特征的表达能力。具体地,参见图5所示,该流量识别模型还包括注意力部500。该注意力部500与最后一个残差部210相连,用于对最后一个残差部210的输出数据引入注意力机制,生成上下文向量,并连接上下文向量以及最后一个残差部210的输出数据,生成流量识别模型的输出数据。可以理解,该最后一个残差部210的输出数据,为该注意力部500的输入数据。

[0110] 本实施例中,在最后一个残差部210的输出侧,增加了注意力部500,利用该注意力部500,通过一系列操作和权重计算,生成一个上下文向量。将上下文向量和之前的输出数据进行连接,以融合注意力信息和其他特征,可以进一步提高对数据的理解能力。

[0111] 可选地,参见图6所示,该注意力部500包括:平均池化层501、重构层502、第二注意力层503、自定义层504、第一点积层505、第三激活层506、第二点积层507、展平层508、第一全连接层509、数据连接层510和第二全连接层511。

[0112] 其中,平均池化层501用于,对最后一个残差部210的输出数据进行平均池化处理。通过平均池化层501对最后一个残差部210的输出数据进行降采样,可以得到一个全局池化的特征向量。其中,该平均池化层501的输出数据具体可以为3D张量的隐藏状态序列(hiddenstates),该张量通常表示来自模型的一系列隐藏状态,其中包括时间步的信息。

[0113] 重构层502用于,对平均池化层501的输出数据进行重构变形。其中,重构层502具体为Reshape层,用于改变输入数据的形状,将数据从一个形状转换为另一个形状。具体地,重构层502可以将一定维度的多维矩阵重新排列构造一个新的保持同样元素数量但是不同维度尺寸的矩阵,实现重构变形。例如,重构层502可以将输入的隐藏状态序列(hiddenstates,隐藏状态在整个模型中是存在的,用于表示输入数据的不同特征和表示)从形状(batch_size,sequence_length,hidden_size)改变为(batch_size,2,2048)。

[0114] 第二注意力层503用于,确定重构层502的输出数据中每个特征的注意力权重。其中,该第二注意力层503与上述第一注意力层3013的原理相似,例如,该第二注意力层503也可以为一个全连接层(Dense)。其中,第一注意力层3013主要用于确定每个特征的重要性,第二注意力层503主要用于针对不同的时间步确定相应的重要性,最终计算得到注意力分数。

[0115] 自定义层504用于,提取重构层502的输出数据中最后一个时间步的隐藏状态。该自定义层504具体可以为Lambda层,该Lambda层可以获取到输入张量(hiddenstates)的最后一个时间步的隐藏状态(一维向量),这个隐藏状态可以被命名为h_t。本实施例中,可以将最后一个时间步的隐藏状态作为当前的查询向量,便于后续计算注意力分数。

[0116] 第一点积层505用于,对第二注意力层503输出的注意力权重与自定义层504输出的隐藏状态进行点积处理(dot),生成注意力分数。

[0117] 第三激活层506用于,基于激活函数对第一点积层505输出的注意力分数进行非线性处理;其中,该第三激活层506与第一激活层3012的工作原理相似,此处不做赘述。例如,该第三激活层506可以使用softmax函数对得到的注意力分数进行归一化,将其处理结果作为一种注意力权重。

[0118] 第二点积层507用于,对第三激活层506的输出数据与自定义层504输出的隐藏状态进行点积处理,生成上下文向量。其中,该第二点积层507与第一点积层505均是用于执行点积处理,只是二者所处理的数据不同。

[0119] 展平层508用于,对平均池化层501的输出数据展平为一维向量。如图6所示,平均池化层501的输出数据除了输入至重构层502,还输入至该展平层508。其中,展平层508具体可以为Flatten层,用于将展平层508的输入数据展平为一维向量。例如,展平层508可以将来自基础模型的输出从二维形状展平为一维形状,以便后续的全连接层处理。

[0120] 第一全连接层509用于,对展平层508的输出数据进行特征提取。其中,该第一全连接层509的数量为多个;如图6所示,可以设有两个串联的第一全连接层509。本实施例中,通过两个全连接层(两个第一全连接层509),可以对特征进行非线性变换和提取,能够在特征表示的基础上进一步提取和加工特征。

[0121] 数据连接层510用于,连接第二点积层507输出的上下文向量以及第一全连接层509的输出数据。该数据连接层510具体可以为Concatenate层,基于可选的轴参数指定要连接的维度,从而可以在指定的轴上连接两个或多个张量。本实施例中,数据连接层510用于连接注意力加权的上下文向量和全连接层的输出;例如,可以将上下文向量与全连接层的输出进行拼接。其中,全连接层的输出本质上可以表示残差部210的输出,即残差网络的输出,将其拼接上下文向量,可以引入注意力机制,以增强模型对输入数据的关注,并提高模型在特定任务上的性能。

[0122] 第二全连接层511用于,对数据连接层510的输出数据进行特征提取,输出流量识别模型的输出数据。本实施例中,通过第二全连接层511将融合后的特征进一步处理,可以输出流量识别模型最终的分类结果,以确定网络流量是否为攻击流量。

[0123] 本实施例中,卷积层202、支路卷积层302、第一卷积层3011以及第二卷积层4011本质上均是一种卷积层,只是其中使用的卷积核大小和/或数量不同。例如,卷积层202采用7×7的卷积核,而支路卷积层302、第一卷积层3011以及第二卷积层4011均采用3×3的卷积核。

[0124] 本实施例提供的攻击流量的识别方法,通过引入注意力机制,通过一系列操作和权重计算,生成上下文向量,将上下文向量和之前的模型输出进行连接,以融合注意力信息和其他特征,使模型能够更智能地处理输入数据,并提高模型在各种任务上的性能。这种注意力机制允许模型有选择性地关注输入数据中的不同部分,从而更好地理解和利用输入数据的信息。

[0125] 可选地,为了调整训练中预测标签的权重,可以运用强化自学习的方式来优化模型的性能,确保预测标签样本在少量的情况下,也能得到良好的预测效果;具体地,上述步骤102“根据训练集对残差网络结构的流量识别模型进行训练”可以包括步骤A1至步骤A2。

[0126] 步骤A1,在测试阶段确定预测标签的F1分数。

[0127] 本实施例中,在测试阶段,当对测试数据进行预测时,可以使用训练好的流量识别模型对测试集的测试样本进行推断,得到预测结果。为了评估模型的性能,一般需要进行分类报告的生成。该分类报告是基于真实标签和预测结果生成的一种评估指标,里面包含了每个预测标签的F1分数。其中,该F1分数具体可基于流量识别模型的精确率和召回率,此处不做详述。

[0128] 步骤A2,根据F1分数对流量识别模型的权重进行更新,且更新后的权重满足:$w_{new} = w_{old} + \frac{1}{F_1 + a}$;其中,$w_{new}$表示更新后的权重,$w_{old}$表示更新前的原始权重,$F_1$表示F1分数,a为防止除零的极小值,例如,0<a<0.1。

[0129] 本实施例中,在得到预测标签的F1分数后,即可对流量识别模型的权重进行更新。具体地,在每个循环迭代中,可以使用当前模型在测试集上的预测结果计算每个类别的F1分数,并根据F1分数的倒数来调整各类别的权重,即 $w_{new} = w_{old} + \frac{1}{F_1 + a}$,将权重从$w_{old}$更新至$w_{new}$,从而增加分类效果较差的类别的权重,实现动态调整类别权重的训练过程。之后,使用调整后的权重进行模型训练,以便更加关注容易被错分的类别,从而提高模型性能。

[0130] 本实施例提供的攻击流量的识别方法,在卷积块中嵌入了注意力机制,增强了卷积块的表征能力,提高了模型的性能,使之更具灵活性和适应性,能够捕捉到从正常流量模式中脱颖而出的异常行为,提高了检测的准确性和覆盖范围,同时降低了误报率。并且,通过调整训练中预测标签类别的权重,可以保证预测标签在少量样本的情况下,也能预测良好,能够优化模型的性能。该方法还能够自动学习网络的正常行为,并实时监测并警报任何异常情况;基于该流量识别模型,可以分析各种网络数据,如网络流量、日志记录、系统事件等,以识别潜在的入侵行为;通过不断学习和适应,能够及时发现新型攻击和未知的安全威胁。

[0131] 上文详细描述了本发明实施例提供的攻击流量的识别方法,该方法也可以通过相应的装置实现,下面详细描述本发明实施例提供的攻击流量的识别装置。

[0132] 图7示出了本发明实施例所提供的一种攻击流量的识别装置的结构示意图。如图7所示,该攻击流量的识别装置包括:

[0133] 获取模块701,用于获取带标签的多个网络样本流量,形成包括多个所述网络样本流量的训练集;

[0134] 训练模块702,用于根据所述训练集对残差网络结构的流量识别模型进行训练;所述流量识别模型嵌有注意力机制;

[0135] 处理模块703,用于生成训练后的流量识别模型;所述训练后的流量识别模型用于识别网络流量中的攻击流量。

[0136] 在一些可选的实施方式中,所述流量识别模型包括依次串接的至少一个残差部;

[0137] 所述残差部包括依次串接的一个卷积块以及至少一个恒等块;所述卷积块嵌有注意力机制。

[0138] 在一些可选的实施方式中,所述卷积块包括:位于主路依次串接的至少一个注意力残差单元,位于支路的支路卷积层,以及第一相加层;所述注意力残差单元包括:第一卷积层、第一激活层、第一注意力层和相乘层;

[0139] 所述支路卷积层用于,对所述卷积块的输入数据进行卷积处理,生成不同通道数的输出数据;

[0140] 所述第一卷积层用于,对所述注意力残差单元的输入数据进行卷积处理;

[0141] 所述第一激活层用于,基于激活函数对所述第一卷积层的输出数据进行非线性处

理；

[0142] 所述第一注意力层用于，确定所述第一激活层的输出数据中每个特征的注意力权重；

[0143] 所述相乘层用于，对所述第一注意力层输出的注意力权重与所述第一激活层的输出数据进行相乘处理；

[0144] 所述第一相加层用于，对所述主路的输出数据与所述支路的输出数据进行相加处理。

[0145] 在一些可选的实施方式中，所述恒等块包括：位于主路依次串接的至少一个恒等残差单元，以及第二相加层；所述恒等残差单元包括：第二卷积层和第二激活层；

[0146] 所述第二卷积层用于，对所述恒等残差单元的输入数据进行卷积处理；

[0147] 所述第二激活层用于，基于激活函数对所述第二卷积层的输出数据进行非线性处理；

[0148] 所述第二相加层用于，对所述主路的输出数据与所述恒等块的输入数据进行相加处理。

[0149] 在一些可选的实施方式中，所述流量识别模型还包括注意力部；

[0150] 所述注意力部与最后一个残差部相连，用于对所述最后一个残差部的输出数据引入注意力机制，生成上下文向量，并连接所述上下文向量以及所述最后一个残差部的输出数据，生成所述流量识别模型的输出数据。

[0151] 在一些可选的实施方式中，所述注意力部包括：平均池化层、重构层、第二注意力层、自定义层、第一点积层、第三激活层、第二点积层、展平层、第一全连接层、数据链接层和第二全连接层；

[0152] 所述平均池化层用于，对所述最后一个残差部的输出数据进行平均池化处理；

[0153] 所述重构层用于，对所述平均池化层的输出数据进行重构变形；

[0154] 所述第二注意力层用于，确定所述重构层的输出数据中每个特征的注意力权重；

[0155] 所述自定义层用于，提取所述重构层的输出数据中最后一个时间步的隐藏状态；

[0156] 所述第一点积层用于，对所述第二注意力层输出的注意力权重与所述自定义层输出的隐藏状态进行点积处理，生成注意力分数；

[0157] 所述第三激活层用于，基于激活函数对所述第一点积层输出的注意力分数进行非线性处理；

[0158] 所述第二点积层用于，对所述第三激活层的输出数据与所述自定义层输出的隐藏状态进行点积处理，生成上下文向量；

[0159] 所述展平层用于，对所述平均池化层的输出数据展平为一维向量；

[0160] 所述第一全连接层用于，对所述展平层的输出数据进行特征提取；

[0161] 所述数据链接层用于，连接所述第二点积层输出的上下文向量以及所述第一全连接层的输出数据；

[0162] 所述第二全连接层用于，对所述数据链接层的输出数据进行特征提取，输出所述流量识别模型的输出数据。

[0163] 在一些可选的实施方式中，所述训练模块702根据所述训练集对残差网络结构的流量识别模型进行训练，包括：

[0164] 在测试阶段确定预测标签的F1分数；

[0165] 根据所述F1分数对所述流量识别模型的权重进行更新，且更新后的权重满足：

[0166] $$w_{new} = w_{old} + \frac{1}{F_1 + a};$$

[0167] 其中，w_{new}表示更新后的权重，w_{old}表示更新前的原始权重，F_1表示所述F1分数，a为防止除零的极小值。

[0168] 本发明实施例还提供了一种计算机存储介质，所述计算机存储介质存储有计算机可执行指令，其包含用于执行上述的全局控制权限的方法的程序，该计算机可执行指令可执行上述任意方法实施例中的方法。

[0169] 其中，所述计算机存储介质可以是计算机能够存取的任何可用介质或数据存储设备，包括但不限于磁性存储器(例如软盘、硬盘、磁带、磁光盘(MO)等)、光学存储器(例如CD、DVD、BD、HVD等)、以及半导体存储器(例如ROM、EPROM、EEPROM、非易失性存储器(NAND FLASH)、固态硬盘(SSD))等。

[0170] 图8示出了本发明的另一个实施例的一种电子设备的结构框图。所述电子设备1100可以是具备计算能力的主机服务器、个人计算机PC、或者可携带的便携式计算机或终端等。本发明具体实施例并不对电子设备的具体实现做限定。

[0171] 该电子设备1100包括至少一个处理器(processor)1110、通信接口(Communications Interface)1120、存储器(memory array)1130和总线1140。其中，处理器1110、通信接口1120、以及存储器1130通过总线1140完成相互间的通信。

[0172] 通信接口1120用于与网元通信，其中网元包括例如虚拟机管理中心、共享存储等。

[0173] 处理器1110用于执行程序。处理器1110可能是一个中央处理器CPU，或者是专用集成电路ASIC(Application Specific Integrated Circuit)，或者是被配置成实施本发明实施例的一个或多个集成电路。

[0174] 存储器1130用于可执行的指令。存储器1130可能包含高速RAM存储器，也可能还包括非易失性存储器(non-volatile memory)，例如至少一个磁盘存储器。存储器1130也可以是存储器阵列。存储器1130还可能被分块，并且所述块可按一定的规则组合成虚拟卷。存储器1130存储的指令可被处理器1110执行，以使处理器1110能够执行上述任意方法实施例中的全局控制权限的方法。

[0175] 本发明实施例通过流程图和/或方框图描述所提供的方法、装置、电子设备。

[0176] 应当理解，流程图和/或方框图的每个方框以及流程图和/或方框图中各方框的组合，都可以由计算机可读程序指令实现。这些计算机可读程序指令可以提供给通用计算机、专用计算机或其他可编程数据处理装置的处理器，从而生产出一种机器，这些计算机可读程序指令通过计算机或其他可编程数据处理装置执行，产生了实现流程图和/或方框图中的方框规定的功能/操作的装置。

[0177] 也可以将这些计算机可读程序指令存储在能使得计算机或其他可编程数据处理装置以特定方式工作的计算机可读存储介质中。这样，存储在计算机可读存储介质中的指令就产生出一个包括实现流程图和/或方框图中的方框规定的功能/操作的指令装置产品。

[0178] 也可以将计算机可读程序指令加载到计算机、其他可编程数据处理装置或其他设备上，使得在计算机、其他可编程数据处理装置或其他设备上执行一系列操作步骤，以产生

计算机实现的过程,从而使得在计算机或其他可编程数据处理装置上执行的指令能够提供实现流程图和/或方框图中的方框规定的功能/操作的过程。

[0179] 以上所述,仅为本发明实施例的具体实施方式,但本发明实施例的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明实施例披露的技术范围内,可轻易想到变化或替换,都应涵盖在本发明实施例的保护范围之内。因此,本发明实施例的保护范围应以权利要求的保护范围为准。

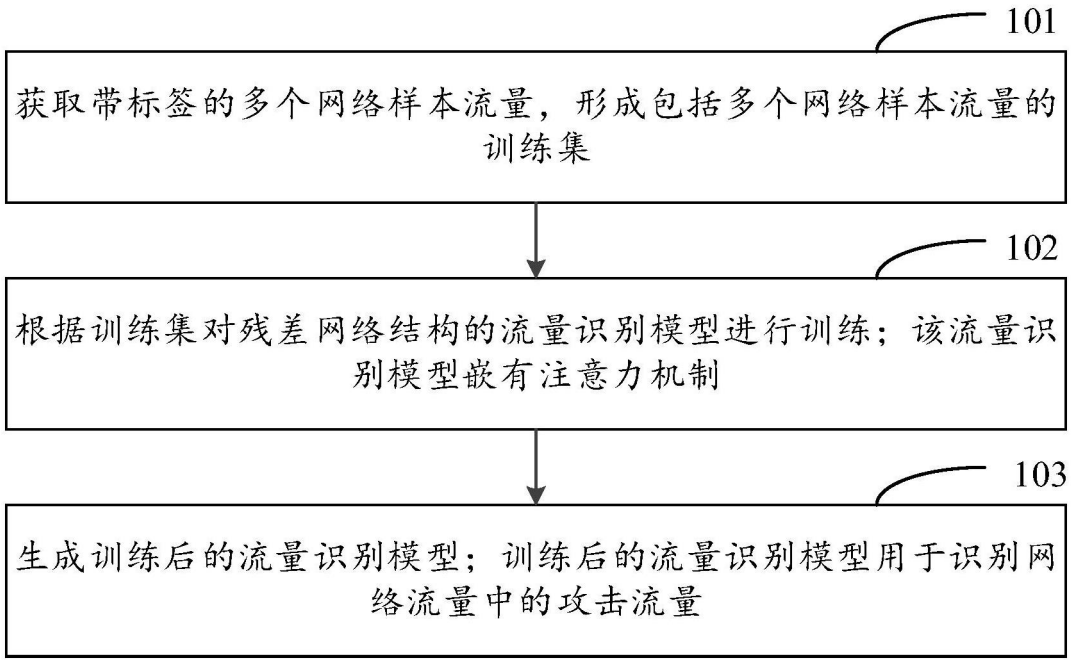

101
获取带标签的多个网络样本流量，形成包括多个网络样本流量的训练集
102
根据训练集对残差网络结构的流量识别模型进行训练；该流量识别模型嵌有注意力机制
103
生成训练后的流量识别模型；训练后的流量识别模型用于识别网络流量中的攻击流量

图1

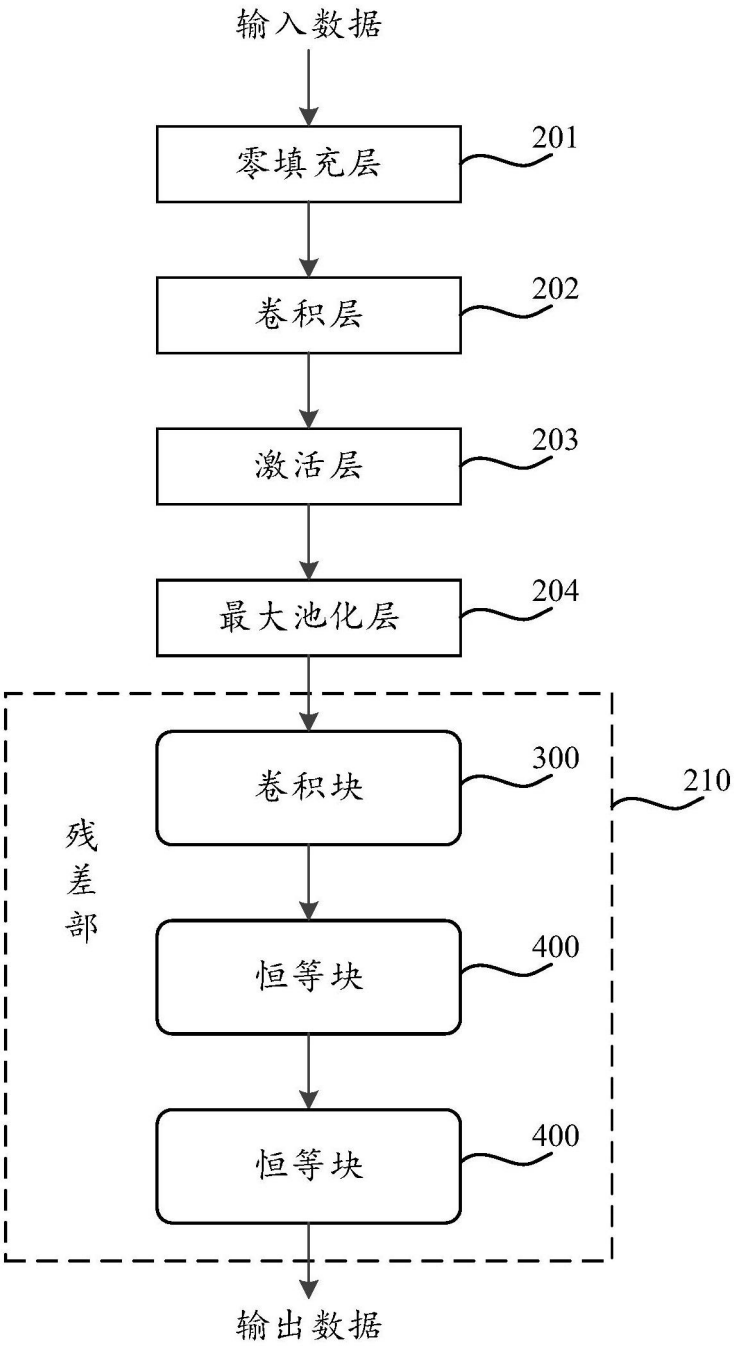

输入数据
零填充层
201
卷积层
202
激活层
203
最大池化层
204
残差部
卷积块
300
210
恒等块
400
恒等块
400
输出数据

图2

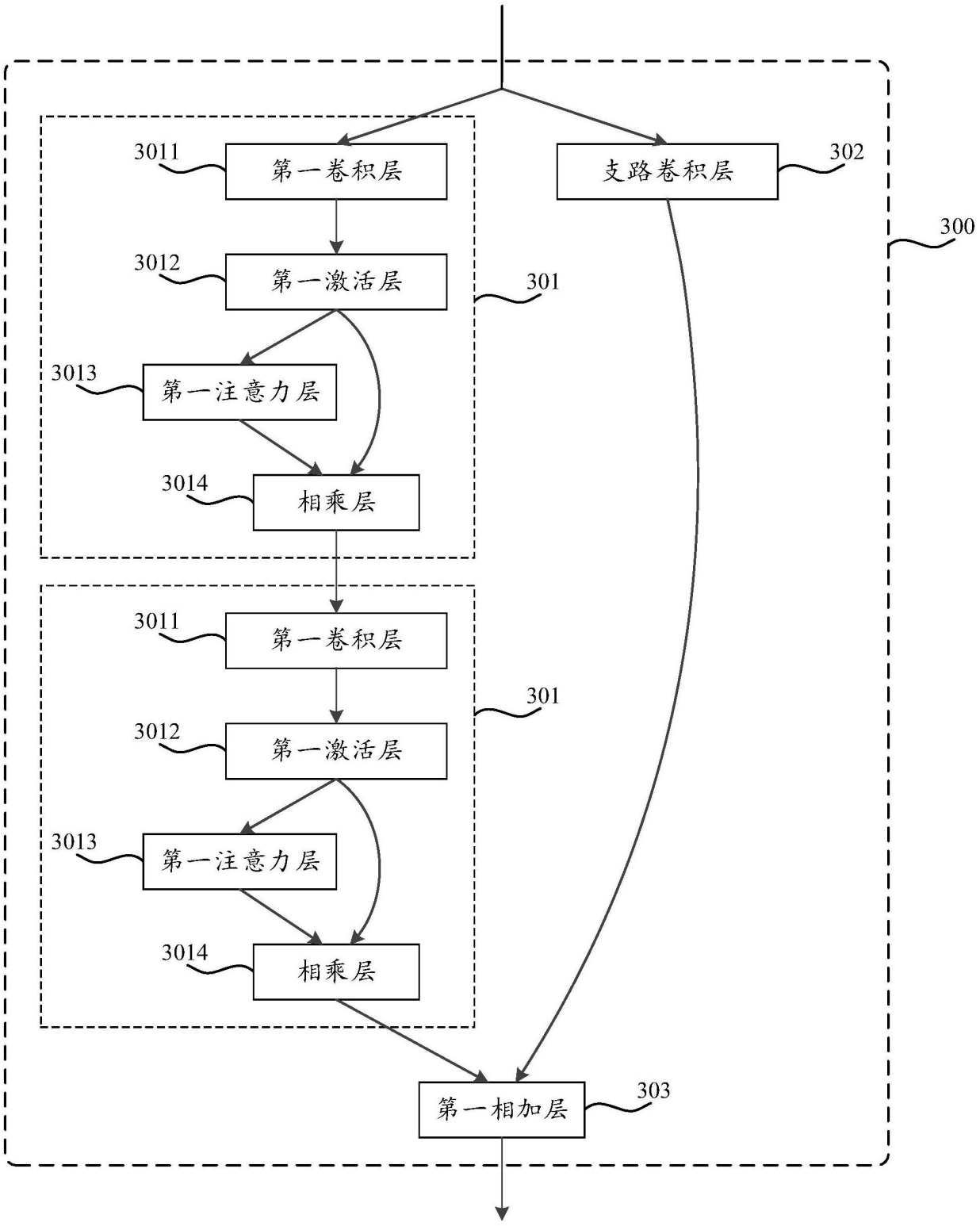

3011
第一卷积层
3012
第一激活层
3013
第一注意力层
3014
相乘层
301
3011
第一卷积层
3012
第一激活层
3013
第一注意力层
3014
相乘层
301
302
支路卷积层
303
第一相加层
300

图3

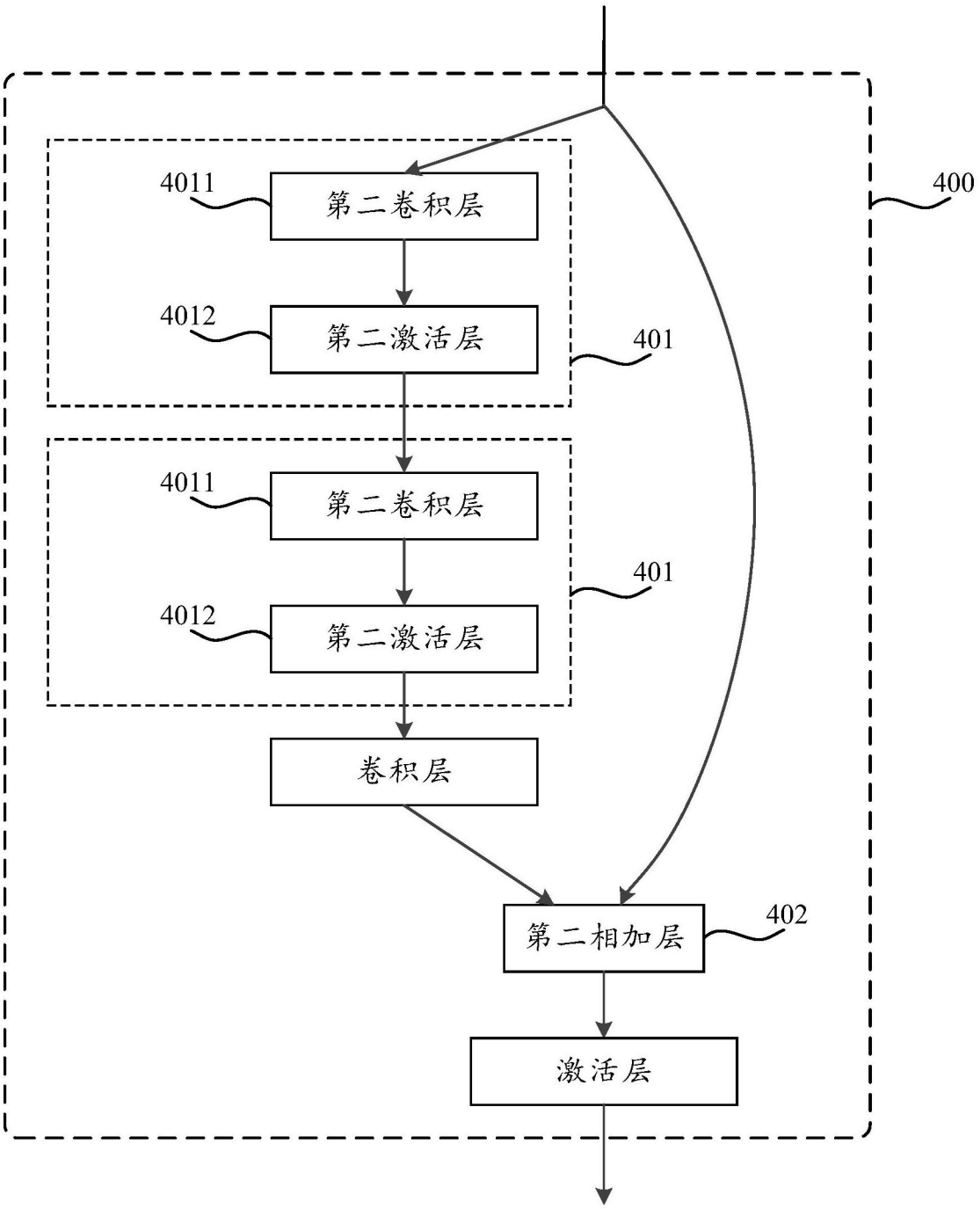

400
4011
第二卷积层
4012
第二激活层
401
4011
第二卷积层
401
4012
第二激活层
卷积层
402
第二相加层
激活层

图4

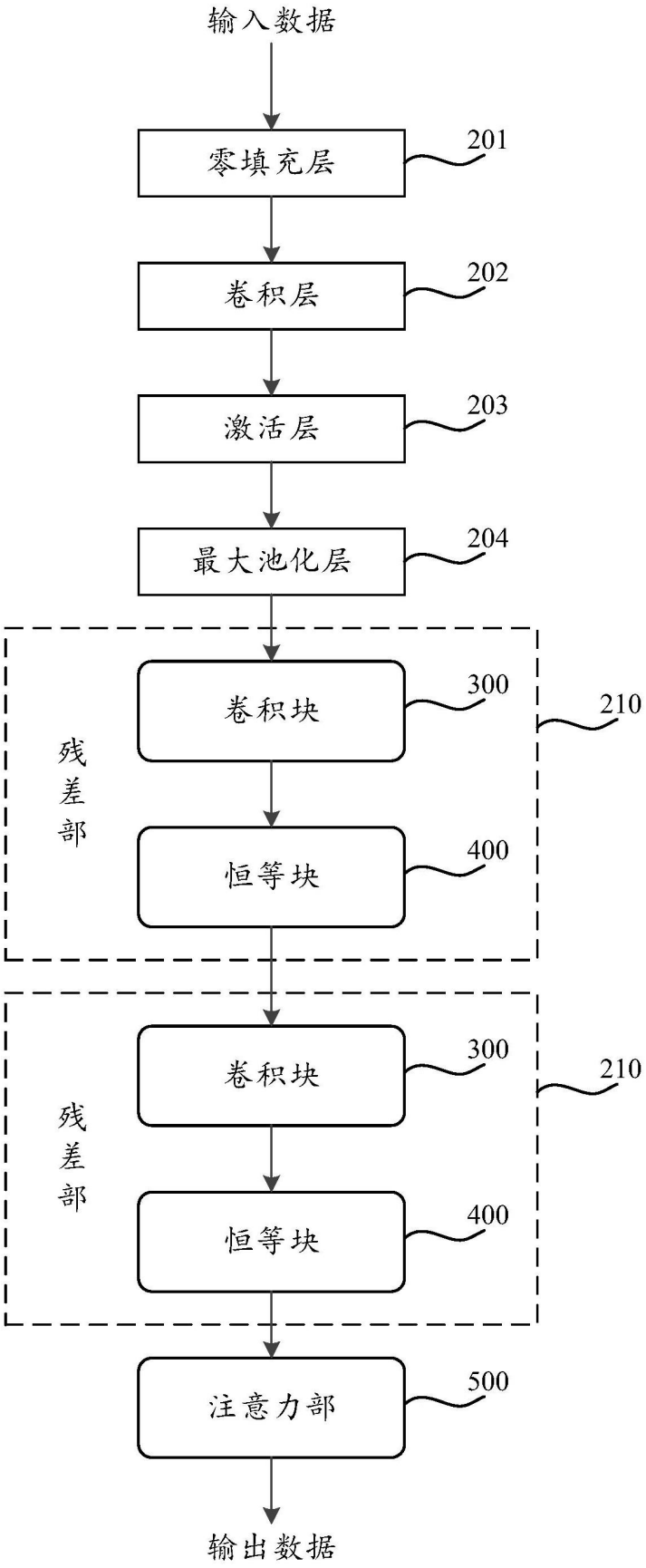

输入数据
零填充层
201
卷积层
202
激活层
203
最大池化层
204
残差部
卷积块
300
210
恒等块
400
残差部
卷积块
300
210
恒等块
400
注意力部
500
输出数据

图5

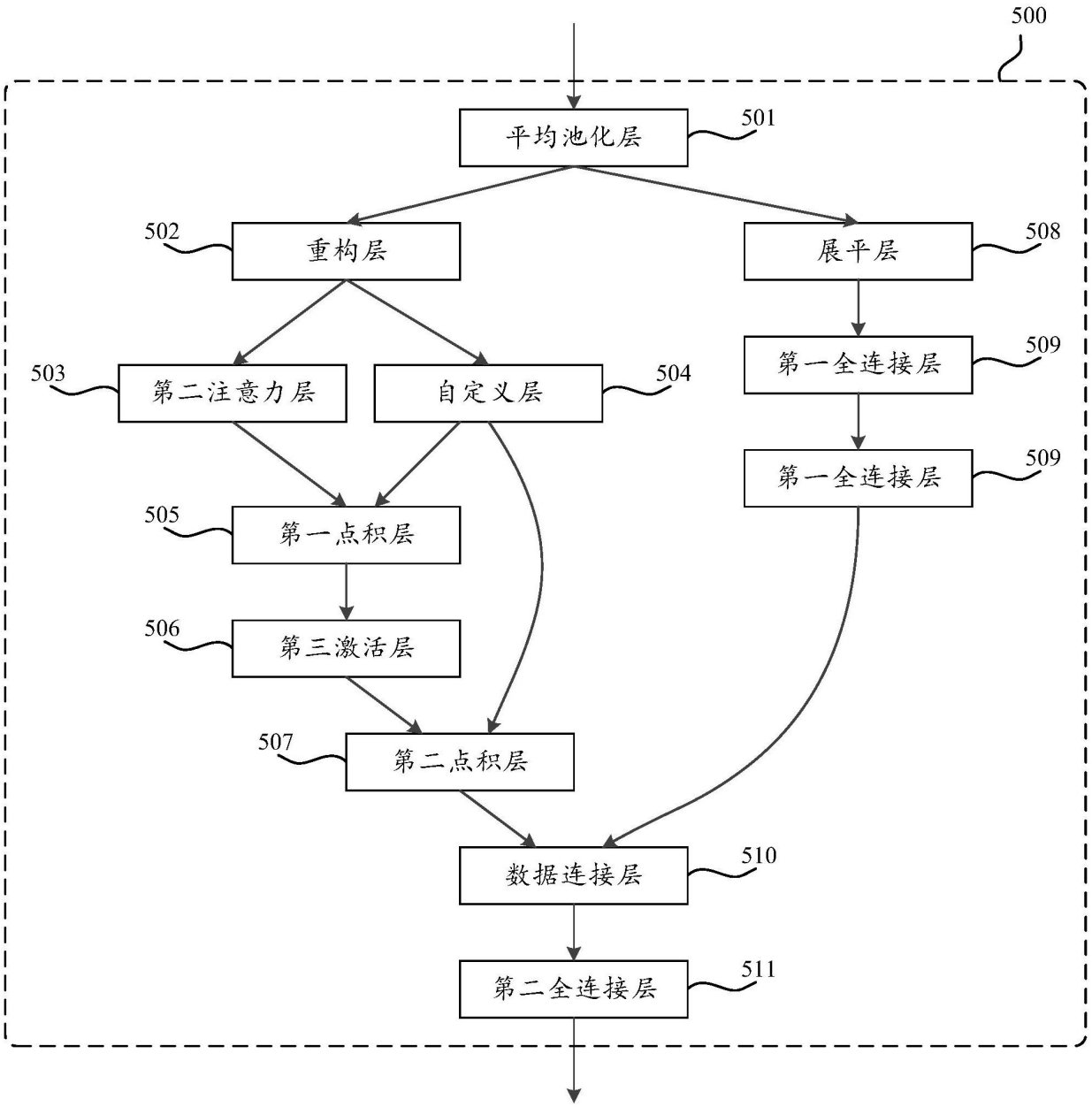

500
平均池化层
501
重构层
502
展平层
508
第二注意力层
503
自定义层
504
第一全连接层
509
第一全连接层
509
第一点积层
505
第三激活层
506
第二点积层
507
数据连接层
510
第二全连接层
511

图6

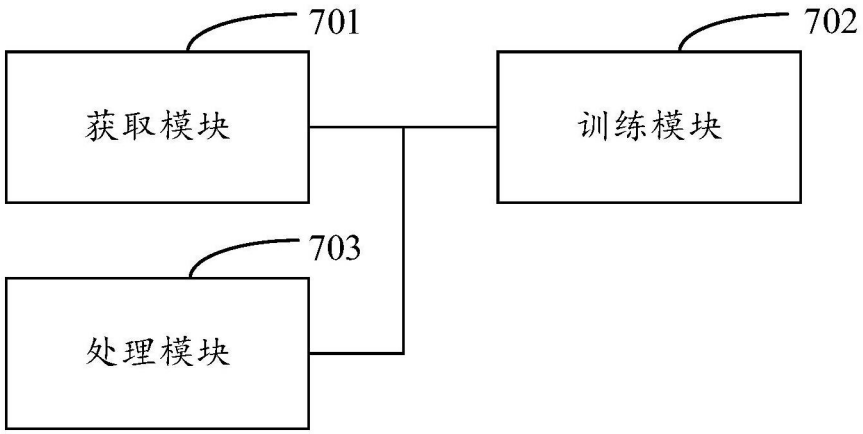

701
获取模块
702
训练模块
703
处理模块

图7

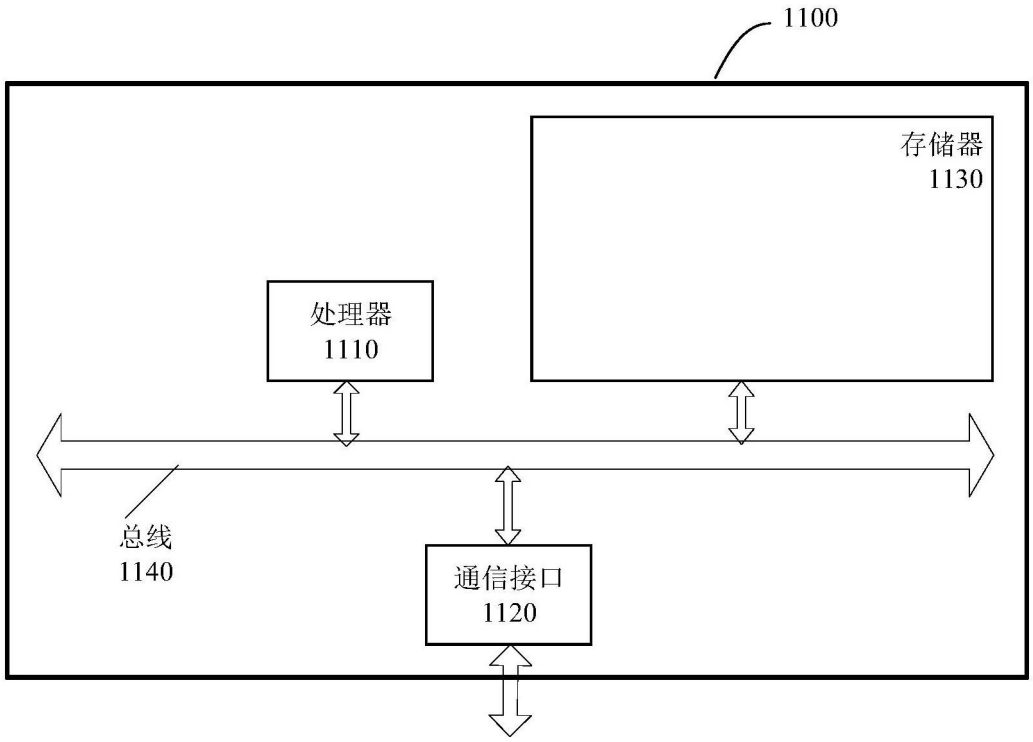

1100
存储器
1130
处理器
1110
总线
1140
通信接口
1120

图8