

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2021 年 7 月 1 日 (01.07.2021)



(10) 国际公布号
WO 2021/129236 A1

- (51) 国际专利分类号:
G06K 9/62 (2006.01) *G01N 29/04* (2006.01)
G06K 9/00 (2006.01)
- (21) 国际申请号: PCT/CN2020/129658
- (22) 国际申请日: 2020 年 11 月 18 日 (18.11.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201911347540.6 2019 年 12 月 24 日 (24.12.2019) CN
- (71) 申请人: 浙江大学(ZHEJIANG UNIVERSITY) [CN/CN]; 中国浙江省杭州市西湖区余杭塘路 866 号, Zhejiang 310058 (CN)。
- (72) 发明人: 邹俊(ZOU, Jun); 中国浙江省杭州市西湖区余杭塘路 866 号, Zhejiang 310058 (CN)。 陆刚(LU, Gang); 中国浙江省杭州市西湖区余杭塘路 866 号, Zhejiang 310058 (CN)。 王渊彬(WANG, Yuanbin); 中国浙江省杭州市西湖区余杭塘路 866 号, Zhejiang 310058 (CN)。

- (74) 代理人: 杭州求是专利事务所有限公司 (HANGZHOU QIUSHI PATENT OFFICE CO., LTD.); 中国浙江省杭州市西湖区玉古路 147 号 黄鸿年科技综合大楼 215 室 / 林超, Zhejiang 310013 (CN)。
- (81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。
- (84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,

(54) Title: ACOUSTIC GARBAGE CLASSIFICATION METHOD USING ONE-DIMENSIONAL CONVOLUTIONAL NEURAL NETWORK

(54) 发明名称: 一种采用一维卷积神经网络的声学垃圾分类方法

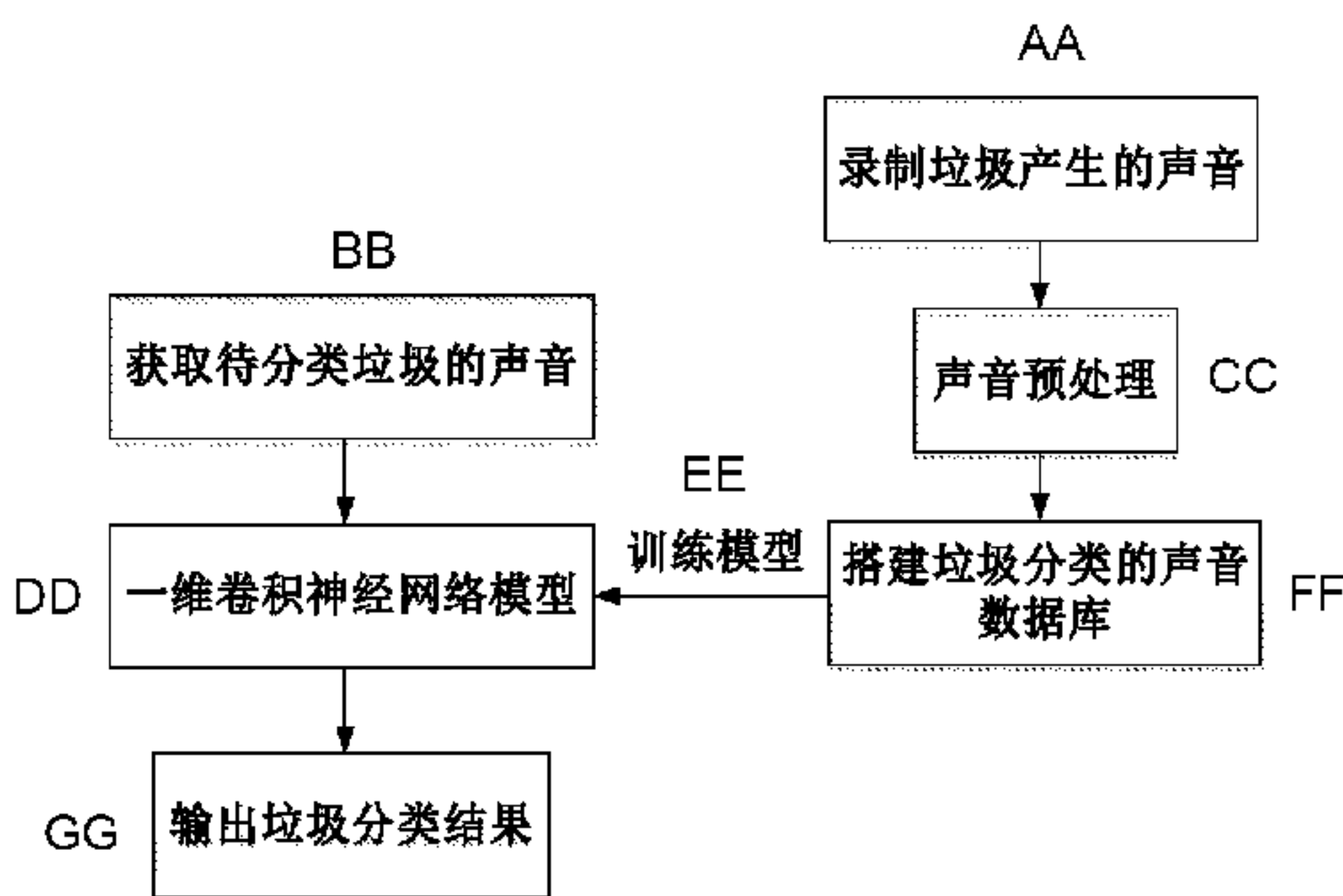


图 1

- AA Record sounds generated by garbage
BB Acquire a sound of garbage to be classified
CC Preprocess the sounds
DD One-dimensional convolutional neural network model
EE Train a model
FF Build a sound database for garbage classification
GG Output a garbage classification result

(57) Abstract: An acoustic garbage classification method using a one-dimensional convolutional neural network. The method comprises: collecting a sound signal generated by the falling of garbage; preprocessing the sound signal; collecting and preprocessing sound signals of different types of garbage to build a sound database for garbage classification, and building and training a one-dimensional convolutional neural network model; and acquiring a sound of garbage to be classified, inputting the sound into the trained one-dimensional convolutional neural network for garbage classification, and outputting a classification result. The garbage classification method is conducive to assisting people in accurately classifying garbage, thereby improving the accuracy of garbage classification and recovery, and the method has extremely high practical and promotional value.

(57) 摘要: 一种采用一维卷积神经网络的声学垃圾分类方法。采集垃圾掉落产生的声音信号; 声音信号预处理; 采集不同种类垃圾的声音信号并预处理搭建垃圾分类的声音数据库, 建立并训练一维卷积神经网络模型; 获取待分类垃圾的声, 输入训练后的一维卷积神经网络进行垃圾分类并输出分类结果。垃圾分类方法有助于辅助人们进行准确地垃圾分类, 从而提高垃圾分类回收的准确率, 具有很高的实用和推广价值。

IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布：

— 包括国际检索报告(条约第21条(3))。

一种采用一维卷积神经网络的声学垃圾分类方法

技术领域

本发明涉及垃圾分类方法，具体地讲，尤其是涉及了一种采用一维卷积神经网络的声学垃圾分类方法。

背景技术

由于人口、城市化和经济的快速发展，城市固体垃圾的数量在急剧增加，这给城市固体垃圾管理带来了巨大的挑战。城市固体垃圾管理包括垃圾的产生、源头分离、储存、收集、运输和处理等。垃圾源头分离是至关重要的一步，因为分类的质量将直接导致这些垃圾后续被如何处理。当前，大多数自动分类技术主要用于垃圾处理厂的批量、大规模分类处理。有效的垃圾源头分离技术仍然是缺乏的。

目前，垃圾的源头分离主要依靠人们的意识和具有的垃圾分类知识，这就需要花费大量的时间和金钱在教育 and 监督上。在许多发展中国家，让人们在短时间内形成垃圾分类的习惯是不现实的。因此，开发一种垃圾源头分离技术在人们的日常生活中帮助他们进行垃圾分类迫在眉睫。图像识别技术已经被广泛的运用在垃圾分类中。然而每一种垃圾可以有各种各样的表现形式，如不同的颜色、形状、变形、与其他物质的混合等。这给图像分类模型训练准备足够的、无偏倚的数据集带来了很大的挑战。物体的声音与它的图像不同，声音与材料属性高度相关，使用声音对垃圾进行分类是一个很好的选择。

如今，涉及声音分类的方法主要有两种。一种方法是人工提取声音数据特征，使用浅层分类器进行分类。这种方法过程繁琐，分类准确率低。另一种方法是将一维的声音数据转换为二维图像，使用二维卷积神经网络进行分类。这种方法计算复杂，导致实时性差。因此，基于一维卷积神经网络的声学垃圾分类方法将快速、有效、实时地对垃圾进行分类，提高垃圾回收率，摆脱垃圾围城困境。

发明内容

本发明的目的在于提供一种采用一维卷积神经网络的声学垃圾分类方法，利用声音进行垃圾的识别和分类，主要解决垃圾源头分离过程中人为垃圾分类困难的问题，从而实现高准确率的垃圾源头分类。

如图 1 所示，本发明采用的技术方案如下：

(A)采集垃圾掉落产生的声音信号；

(B)声音信号预处理；

(C)采集不同种类垃圾的声音信号并预处理搭建垃圾分类的声音数据库，建立并训练一维卷积神经网络模型；

(D)获取待分类垃圾的声，输入训练后的一维卷积神经网络进行垃圾分类并输出分类结果。

本发明基于声音进行深度学习处理进而实现垃圾分类的方法有助于准确地垃圾分类，从而提高垃圾分类回收的准确率，具有很高的实用和推广价值。

所述步骤(A)中，垃圾掉落产生的声音是将垃圾自由落体撞击到一块板上产生的，产生的撞击声被一个单通道麦克风记录；采样频率为 44100Hz；每种垃圾重复上述过程多次获得多段声音信号。

所述步骤(B)中，预处理为：每段声音信号只截取 120ms 有效时长，120ms 时长的声音信号以每段声音信号的峰值为起点向后 120ms 为终点，，这样使得这段声音信号覆盖整个撞击过程，并能去除声音信号中的冗余信息。

所述步骤(C)中，一维卷积神经网络模型主要包括输入层、卷积层、池化层、全连接层和输出层；网络模型训练时，输入层为已标注垃圾类型的预处理后的声音信号；卷积层对上一层的输出数据进行卷积运算、提取特征，采用 ReLU 激活函数以增强卷积神经网络的非线性性能，接着经过最大池化层进行特征降维、减少网络参数数量和过拟合，最后通过全连接层、输出层进行分类并输出分类结果。

所述步骤(C)中，一维卷积神经网络模型训练过程中，不同的超参数的选择将影响网络的分类性能；本发明主要考虑三个超参数：网络深度、卷积核大小和学习率；考虑到训练时间、计算复杂度和计算机配置问题，每个参数只选取五个经验值；采用正交实验法获得三个参数值的最优组合，使得一维卷积神经网络模型的分类性能最优。

本发明具有以下有益效果：

(1)本发明创造性地使用声音数据进行垃圾分类，声音与材料属性高度相关，声音信号可以揭示垃圾的固有特性，如弹性和内摩擦。弹性影响产生声音的频率。内摩擦决定了产生的声音如何随着时间衰减，并为垃圾分类提供了形状不同的声学特征。同时，使用声音进行垃圾分类大大减少了训练模型的数据集。

(2)本发明在一维卷积神经网络训练过程中，采用正交实验法研究不同超参数对一维卷积神经网络分类性能的影响，最终确定最优的超参数组合，从而使网络模型获得最优分类性能。

(3)本发明通过一维卷积神经网络进行垃圾分类，一维卷积神经网络适用于处理一维信号，能够直接从原始数据中提取到最优特征，避免了人工提取特征的繁琐性、次优性以及二维卷积神经网络分类的计算复杂性，最大限度地提高了分类准确度。同时，一维卷积神经网络适合于可移动、低功耗、低内存设备上的实时应用。

附图说明

图 1 为本发明的方法流程图；

图 2 为基于声音信号的垃圾分类卷积神经网络模型图。

图 3 为本发明的一维卷积神经网络的结构示意图。

具体实施方式

下面结合附图和实施例对本发明作进一步说明，本发明的实施方式包括但不限于下列实施例。

本发明的具体实施例及其实施情况如下：

具体实施的一维卷积神经网络模型选择 TensorFlow 1.9 作为深度学习框架，所有的算法用 python 3.6 编写，所有程序在一台带有 GTX1050Ti 显卡、第 8 代 Intel Core i5 处理器和 512GB SSD 的笔记本电脑上执行。

如图 2 所示，基于声音采集装置，准备各类垃圾材料，然后进行垃圾声音的采集与标注，对这些数据进行预处理，设计一种简单高效的网络结构来分析声音数据。卷积网络中包含了大量参数，采用正交实验法，搜索最优的参数组合；在正交实验过程中，网络深度分别取 1、2、3、4、5，卷积核大小分别取 3、5、7、9、11，学习率分别取 $1e-5$ 、 $5e-5$ 、 $1e-4$ 、 $5e-4$ 、 $1e-3$ ，依据六因素五水平的正交设计表进行 25 次实验，最终获得最优的参数组合为：网络深度取 5，卷积核大小取 9，学习率取 $5e-4$ 。

如图 3 所示，实施例的一维卷积神经网络模型主要包括输入层、五个卷积层、五个池化层、全连接层和输出层。一维声音信号被输入到一维卷积神经网络的输入层。在输入信号和对应的卷积核之间进行卷积运算，生成输入特征映射。然后通过激活函数传递输入特征映射，生成卷积层的输出特征映射。在卷积层之后，通常采用池化层，不仅可通过降低从卷积层提取的特征维数来减少计算成本，而且对特征提供了基本的平移不变性。

实施例的一维卷积神经网络模型的具体设计为：定义卷积层 1-5，卷积核个数分别为 8，16，24，32 和 40，核大小为 9×9 ，步幅 strides 为 1，边界填充 padding 为 SAME，激活函数选用 ReLU 函数；定义最大池化层 1-5，卷积核个数分别为 8，16，24，32 和 40，核大小为 9×9 ，步幅 strides 为 2；最后定义全

连接层和输出层，并使用 softmax 函数进行分类，并输出所属类别的标签 0-n。

训练及评估模型。训练过程主要分为前向传播和反向传播两个阶段。前向传播就是搭建了模型的计算过程，建立了输入与输出的映射关系；反向传播就是使用梯度下降法训练模型参数，实现损失函数最小。获得最终模型后，使用声音数据库中的测试集对模型分类效果进行评估。

(D)获取待分类垃圾的声音，输入训练后的一维卷积神经网络进行垃圾分类并输出分类结果。

以上所述仅为本发明的优先实施方式，本发明并不限于上述实施方式，只要以基本相同手段实现本发明目的的技术方案都属于本发明的保护范围之内。

权 利 要 求 书

1.一种采用一维卷积神经网络的声学垃圾分类方法，其特征在于：方法包括以下步骤：

(A)采集垃圾掉落产生的声音信号；

(B)声音信号预处理；

(C)采集不同种类垃圾的声音信号并预处理搭建垃圾分类的声音数据库，建立并训练一维卷积神经网络模型；

(D)获取待分类垃圾的声，输入训练后的一维卷积神经网络进行垃圾分类并输出分类结果。

2.根据权利要求书 1 所述的一种采用一维卷积神经网络的声学垃圾分类方法，其特征在于：所述步骤(A)中，垃圾掉落产生的声音是将垃圾自由落体撞击到一块板上产生的，产生的撞击声被一个单通道麦克风记录；采样频率为 44100Hz；每种垃圾重复上述过程多次获得多段声音信号。

3.根据权利要求书 1 所述的一种采用一维卷积神经网络的声学垃圾分类方法，其特征在于：所述步骤(B)中，预处理为：每段声音信号只截取 120ms 有效时长，120ms 时长的声音信号以每段声音信号的峰值为起点向后 120ms 为终点。

4. 根据权利要求书 1 所述的一种采用一维卷积神经网络的声学垃圾分类方法，其特征在于：所述步骤(C)中，一维卷积神经网络模型主要包括输入层、卷积层、池化层、全连接层和输出层；网络模型训练时，输入层为已标注垃圾类型的预处理后的声音信号；卷积层对上一层的输出数据进行卷积运算、提取特征，采用 ReLU 激活函数以增强卷积神经网络的非线性性能，接着经过最大池化层进行特征降维、减少网络参数数量和过拟合，最后通过全连接层、输出层进行分类并输出分类结果。

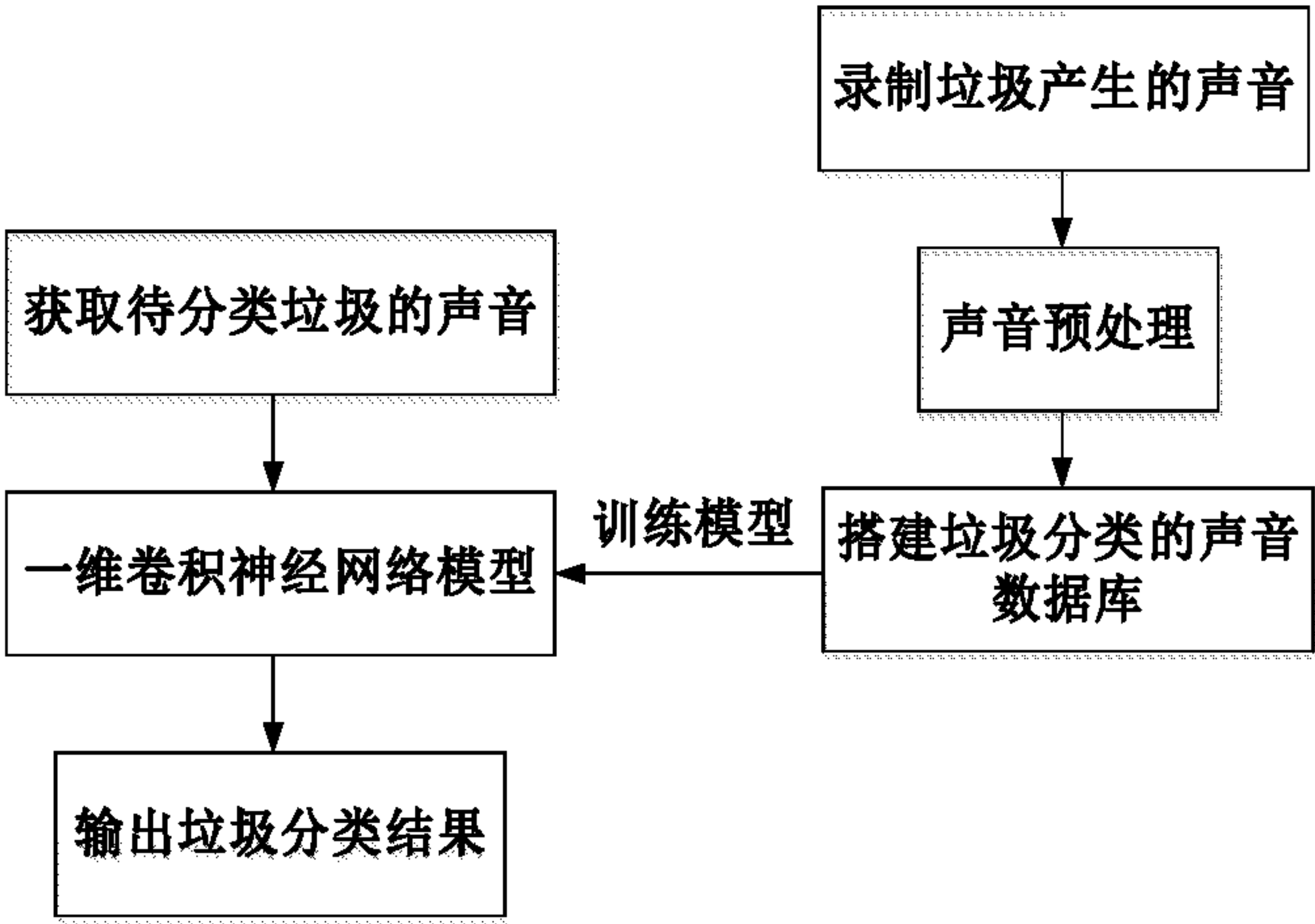


图 1

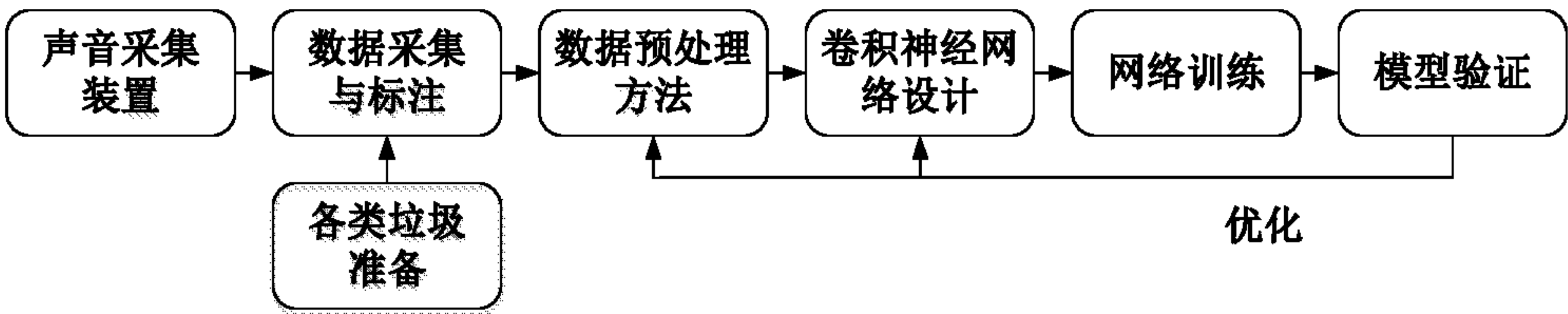


图 2

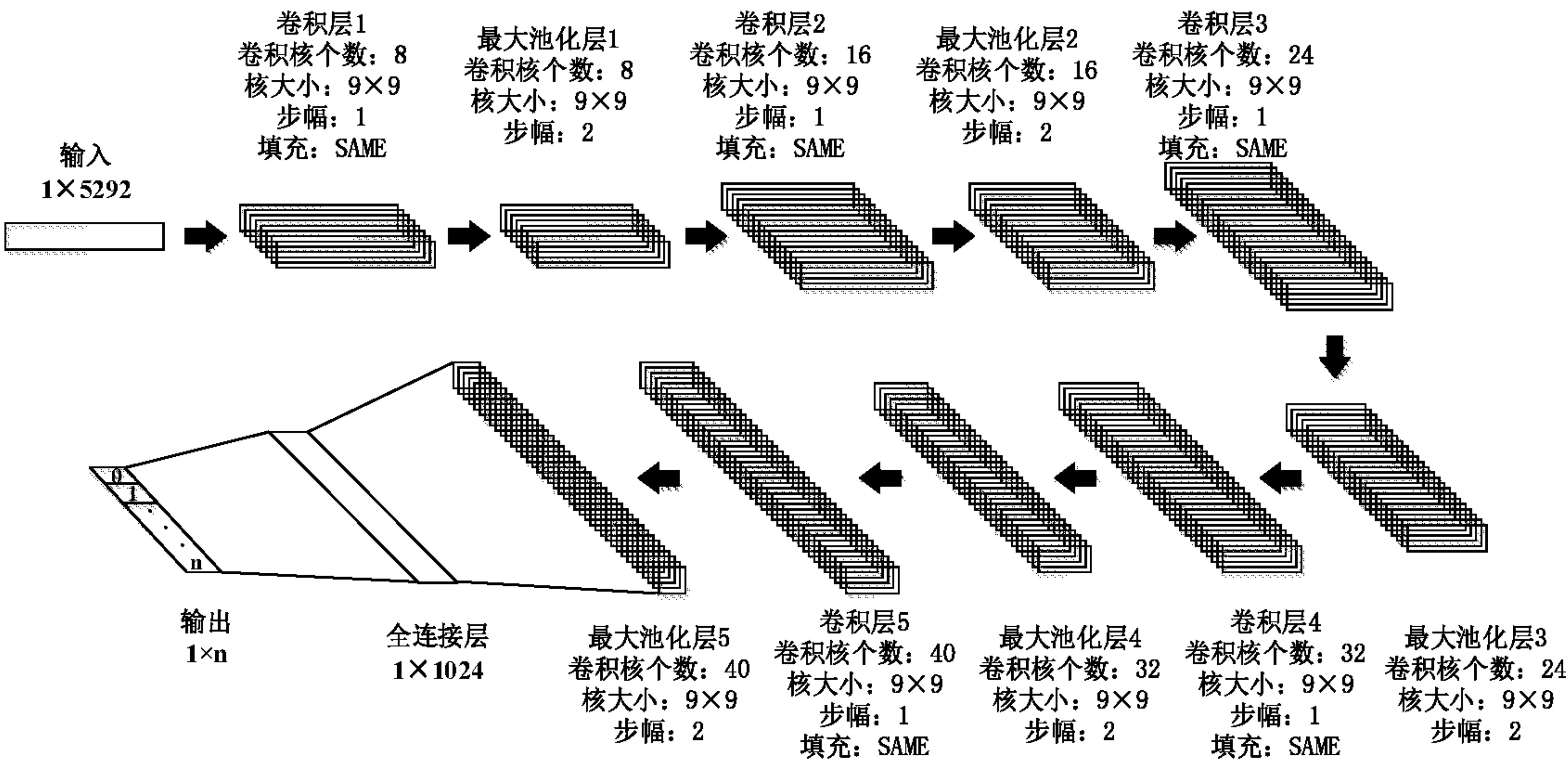


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/129658

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/62(2006.01)i; G06K 9/00(2006.01)i; G01N 29/04(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K; G06F; G06N; B65F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNKI, CNPAT, WPI, EPODOC, IEEE: 垃圾, 分类, 识别, 声学, 声音, 声响, 响声, 信号, 掉落, 丢弃, 撞击, 碰撞, 敲击, 一维, 卷积, 神经网络, garbage, trash, rubbish, waste, classificat+, sorting, recognit+, acoustics, sound, noise, signal, falling, drop+, hit+, knock+, one-dimension, convolution, neural network, model

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 111160438 A (ZHEJIANG UNIVERSITY) 15 May 2020 (2020-05-15) claims 1-4	1-4
PX	CN 110937280 A (ZHEJIANG UNIVERSITY) 31 March 2020 (2020-03-31) description, paragraphs 0005-0027, 0047-0058, figures 1-3	1-4
X	刘强 等 (LIU, Qiang et al.). "基于倒谱与BP网络的船舶生活垃圾分类方法研究 (Research on Classification Methods for Marine Life Rubbish Based on Cepstrum and BP Network)" 《南通航运职业技术学院学报》 (Journal of Nantong Vocational & Technical Shipping College), Vol. 15, No. 1, 31 March 2016 (2016-03-31), ISSN: 1671-9891, main body, sections 1.1, 2, pages 34, 36-38	1-4
A	CN 110427896 A (CHENGDU UNIVERSITY OF TECHNOLOGY) 08 November 2019 (2019-11-08) entire document	1-4
A	KR 20190071387 A (AHN, Sungjun et al.) 24 June 2019 (2019-06-24) entire document	1-4

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance
“E” earlier application or patent but published on or after the international filing date
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
“O” document referring to an oral disclosure, use, exhibition or other means
“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“&” document member of the same patent family

Date of the actual completion of the international search

02 February 2021

Date of mailing of the international search report

19 February 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/129658

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EP 1405706 A1 (NIPPON STEEL CHEMICAL CO., LTD.) 07 April 2004 (2004-04-07) entire document	1-4

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2020/129658

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	111160438	A	15 May 2020	None			
CN	110937280	A	31 March 2020	None			
CN	110427896	A	08 November 2019	None			
KR	20190071387	A	24 June 2019	None			
EP	1405706	A1	07 April 2004	JP	2004122575	A	22 April 2004
				JP	2004122576	A	22 April 2004

国际检索报告		国际申请号 PCT/CN2020/129658
A. 主题的分类 G06K 9/62(2006.01)i; G06K 9/00(2006.01)i; G01N 29/04(2006.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G06K; G06F; G06N; B65F 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) CNKI, CNPAT, WPI, EPODOC, IEEE: 垃圾, 分类, 识别, 声学, 声音, 声响, 响声, 信号, 掉落, 丢弃, 撞击, 碰撞, 敲击, 一维, 卷积, 神经网络, garbage, trash, rubbish, waste, classificat+, sorting, recognit+, acoustics, sound, noise, signal, falling, drop+, hit+, knock+, one-dimension, convolution, neural network, model		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 111160438 A (浙江大学) 2020年 5月 15日 (2020 - 05 - 15) 权利要求1-4	1-4
PX	CN 110937280 A (浙江大学) 2020年 3月 31日 (2020 - 03 - 31) 说明书第0005-0027, 0047-0058段、附图1-3	1-4
X	刘强 等. “基于倒谱与BP网络的船舶生活垃圾分类方法研究” 《南通航运职业技术学院学报》, 第15卷, 第1期, 2016年 3月 31日 (2016 - 03 - 31), ISSN: 1671-9891, 正文第1.1部分、第2部分, 第34、36-38页	1-4
A	CN 110427896 A (成都理工大学) 2019年 11月 8日 (2019 - 11 - 08) 全文	1-4
A	KR 20190071387 A (AHN, Sungjun 等) 2019年 6月 24日 (2019 - 06 - 24) 全文	1-4
A	EP 1405706 A1 (NIPPON STEEL CHEMICAL CO., LTD) 2004年 4月 7日 (2004 - 04 - 07) 全文	1-4
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期 2021年 2月 2日		国际检索报告邮寄日期 2021年 2月 19日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		受权官员 张雯 电话号码 86-(10)-53961314

国际申请号
PCT/CN2020/129658

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	111160438	A	2020年 5月 15日	无	
CN	110937280	A	2020年 3月 31日	无	
CN	110427896	A	2019年 11月 8日	无	
KR	20190071387	A	2019年 6月 24日	无	
EP	1405706	A1	2004年 4月 7日	JP 2004122575 A	2004年 4月 22日
				JP 2004122576 A	2004年 4月 22日