

(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
2020 年 5 月 22 日 (22.05.2020)



(10) 国际公布号
WO 2020/098256 A1

(51) 国际专利分类号:
G10L 21/02 (2013.01) *G10L 25/00* (2013.01)
G10L 21/0208 (2013.01)

(21) 国际申请号: PCT/CN2019/089180

(22) 国际申请日: 2019 年 5 月 30 日 (30.05.2019)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201811350813.8 2018 年 11 月 14 日 (14.11.2018) CN

(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518033 (CN)。

(72) 发明人: 赵峰(ZHAO, Feng); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518033 (CN)。
王健宗(WANG, Jianzong); 中国广东省深圳市

福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518033 (CN)。肖京(XIAO, Jing); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518033 (CN)。

(74) 代理人: 北京鸿元知识产权代理有限公司(GRANDER IP LAW FIRM); 中国北京市朝阳区光华路 7 号汉威大厦东区 18A6 室, Beijing 100004 (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, 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,

(54) Title: SPEECH ENHANCEMENT METHOD BASED ON FULLY CONVOLUTIONAL NEURAL NETWORK, DEVICE, AND STORAGE MEDIUM

(54) 发明名称: 基于全卷积神经网络的语音增强方法、装置及存储介质

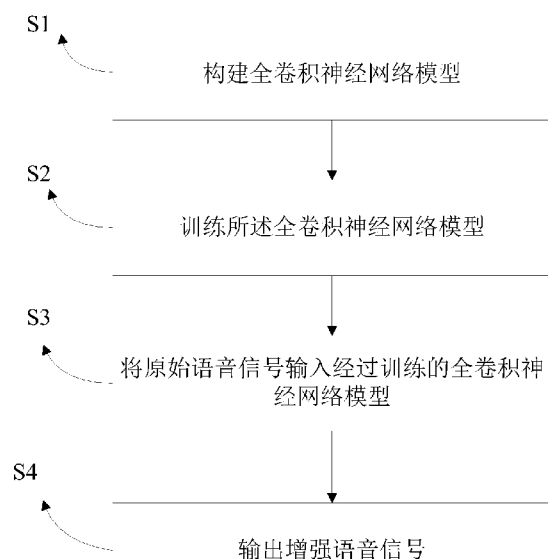


图 1

- S1 Construct a fully convolutional neural network model
- S2 Train the fully convolutional neural network model
- S3 Input an original speech signal into the trained fully convolutional neural network model
- S4 Output an enhanced speech signal

(57) Abstract: The present application relates to the field of artificial intelligence. Disclosed is a speech enhancement method based on a fully convolutional neural network. The method comprises: constructing a fully convolutional neural network model, the fully convolutional neural network comprising an input layer, a hidden layer, and an output layer; the hidden layer comprising a plurality of convolutional layers; each of the plurality of convolutional layers comprising a plurality of filters; training the fully convolutional neural network model; inputting an original speech signal into the trained fully convolutional neural network model; and outputting an enhanced speech signal. In the fully convolutional neural network model of the present application, a full connection layer is deleted, and only convolutional layers are comprised, so that parameters of the neural network are significantly reduced, the fully convolutional neural network model can be suitable for a mobile device having the memory limited, each output sample only relies on adjacent inputs, and original information and spatial arrangement information of the speech signal can be reserved well at less weight values. Also disclosed are an electronic device and a computer readable storage medium.

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, 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,
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))。

(57) 摘要: 本申请属于人工智能领域, 本申请公开了一种基于全卷积神经网络的语音增强方法, 包括: 构建全卷积神经网络模型, 所述全卷积神经网络模型包括输入层、隐含层和输出层, 所述隐含层为多个卷积层, 每个卷积层均具有多个滤波器; 训练所述全卷积神经网络模型; 将原始语音信号输入经过训练的全卷积神经网络模型; 输出增强语音信号。本申请的全卷积神经网络模型中, 删除了全连接层, 仅包含卷积层, 大大减小了神经网络的参数, 使得全卷积神经网络模型可以适用于限制内存的移动设备中, 并且每个输出样本仅仅依赖相邻输入, 可以通过更少的权重值很好地保留语音信号的原始信息和空间排列信息。本申请还公开了一种电子装置和计算机可读存储介质。

基于全卷积神经网络的语音增强方法、装置及存储介质

技术领域

5 本申请涉及语音技术领域，尤其涉及一种基于全卷积神经网络的语音增强方法、装置及存储介质。

背景技术

10 语音增强，是指干净语音在现实的生活场景中受到各种噪声干扰时，需要通过一定的方法将噪声滤除，以提升该段语音的质量和可懂度的技术。日常生活中，麦克风采集的语音通常是带有不同噪声的“污染”语音，语音增强的主要目的就是这些被“污染”的带噪语音中恢复出干净语音。语音增强涉及的应用领域十分广泛，包括语音通话、电话会议、场景录音、军事窃听、助听器设备和语音识别设备等，并成为许多语音编码和识别系统的预处

15 理模块。以语音增强应用于助听器为例，通常的助听器，只是实现一个语音的基本放大，复杂一些的会进行声压级压缩以实现对患者听觉范围的补偿，但是如果听觉场景比较复杂，患者听到的语音中不仅包含了放大后的语音也包含了很多噪声，时间一长势必会对患者的听觉系统造成二次损害，因此高端的数字助听器设备中，语音增强也成为不容忽视的一个重要方面。

20 语音增强应用在语音相关应用的前端处理过程中，确保把语音从带噪信号中分离出来，以便后端识别模型正确识别语音的内容。现有的语音增强方法包括无监督语音增强方法和有监督语音增强方法，其中，无监督语音增强方法是提取语音信号的幅度谱或对数谱，相位信息被忽略，当将语音信号合成到时域时，相位信号不变的应用带噪语音信号的相位信息，会减弱增强语音信号的质量。有监督语音增强方法是基于神经网络的语音增强方法，而采

25 用带有全连接层的深度神经网络（DNN，Deep Neural Network）和卷积神经网络（CNN，Convolutional Neural Network）进行有监督的语音增强，不能很好的表示模型的高低频成分，并且其中的全连接层也不能很好地保留信号的原始信息和空间排列信息。

发明内容

鉴于以上问题，本申请提供一种基于全卷积神经网络的语音增强方法、装置及存储介质，以解决现有的语音增强方法的神经网络模型不能很好地保留语音信号的原始信息和空间排列信息的问题。

5 为了实现上述目的，本申请提供一种基于全卷积神经网络的语音增强方法，包括：

构建全卷积神经网络模型，所述全卷积神经网络模型包括输入层、隐含层和输出层，所述隐含层为多个卷积层，每个卷积层均具有多个滤波器，所述输出层的输出模型为：

$$10 \quad y_t = F^T * R_t \quad (1)$$

其中， y_t 是输出层的第 t 个节点， F^T 是滤波器的权重矩阵的转置， $F \in R^{f \times 1}$ ， f 表示滤波器尺寸， R_t 是隐含层的第 t 个节点；

训练所述全卷积神经网络模型；

将原始语音信号输入经过训练的全卷积神经网络模型；

15 输出增强语音信号。

优选地，根据下式构建全卷积神经网络模型的隐含层的模型：

$$h_j^1 = f(\sum_{i=1}^n (w_{ik}^1 x_i) + a_k^1) \quad (2)$$

$$h_k^l = f(\sum_{j=1}^H (w_{kj}^l h_j^{l-1}) + b_k^l) \quad l \in (2, 3, \dots, L) \quad (3)$$

其中， h_j^1 表示第 1 个隐含层的第 j 个节点的输出值， x_i 表示输入层的第 i 个节点的变量， w_{ik}^1 表示输入层第 i 个节点和第 1 个隐含层第 k 个节点的连接权重值， a_k^1 表示第 1 个隐含层第 k 个节点的偏移量， n 表示输入层的节点数， h_k^l 表示第 l 个隐含层的第 k 个节点的输出值， h_j^{l-1} 表示第 $l-1$ 个隐含层的第 j 个节点的输出值， w_{kj}^l 表示第 l 个隐含层的第 k 个节点和第 $l-1$ 个隐含层的第 j 个节点的连接权重值， b_k^l 表示第 l 个隐含层第 k 个节点的偏移量， H 为隐含层的节点数， f 为激励函数。

25

为了实现上述目的，本申请的另一个方面是提供一种电子装置，该电子装置包括：存储器和处理器，所述存储器中包括语音增强程序，所述语音增强程序被所述处理器执行时实现如下步骤：

构建全卷积神经网络模型，所述全卷积神经网络模型包括输入层、隐含层和输出层，所述隐含层为多个卷积层，每个卷积层均具有多个滤波器，所述输出层的输出模型为：

30

$$y_t = F^T * R_t \quad (1)$$

其中, y_t 是输出层的第 t 个节点, F^T 是滤波器的权重矩阵的转置, $F \in R^{f \times 1}$, f 表示滤波器尺寸, R_t 是隐含层的第 t 个节点;

训练所述全卷积神经网络模型;

- 5 将原始语音信号输入经过训练的全卷积神经网络模型;
输出增强语音信号。

优选地, 所述全卷积神经网络模型中隐含层的模型为:

$$h_j^1 = f(\sum_{i=1}^n (w_{ik}^1 x_i) + a_k^1) \quad (2)$$

$$h_k^l = f(\sum_{j=1}^H (w_{kj}^l h_j^{l-1}) + b_k^l) \quad l \in (2, 3, \dots, L) \quad (3)$$

- 10 其中, h_j^1 表示第 1 个隐含层的第 j 个节点的输出值, x_i 表示输入层的第 i 个节点的变量, w_{ik}^1 表示输入层第 i 个节点和第 1 个隐含层第 k 个节点的连接权重值, a_k^1 表示第 1 个隐含层第 k 个节点的偏移量, n 表示输入层的节点数, h_k^l 表示第 l 个隐含层的第 k 个节点的输出值, h_j^{l-1} 表示第 $l-1$ 个隐含层的第 j 个节点的输出值, w_{kj}^l 表示第 l 个隐含层的第 k 个节点和第 $l-1$ 个隐含层的第 j 个节点的连接权重值, b_k^l 表示第 l 个隐含层第 k 个节点的偏移量, H 为隐含层的节点数, f 为激励函数。

优选地, 处理器训练所述全卷积神经网络模型包括:

对所述全卷积神经网络模型的参数进行初始赋值, 所述参数包括输入层和隐含层的连接权重值、相邻隐含层之间的连接权重值和隐含层的偏移量;

- 20 构建样本集, 并将所述样本集按比例划分为训练样本集和测试样本集;

输入所述训练样本集中的一个训练样本, 并从所述训练样本中提取特征向量;

将训练样本的输入数据代入公式 (1) - (3), 计算隐含层各节点的输出值和输出层各节点的输出值;

- 25 计算输出层各节点误差:

$$e_k = o_k - y_k \quad (4)$$

其中, e_k 表示输出层第 k 个节点的误差, o_k 表示输出层第 k 个节点的实际值, y_k 表示输出层第 k 个节点的输出值;

基于误差反向传播更新所述全卷积神经网络模型的参数;

- 30 输入下一个训练样本, 继续更新全卷积神经网络模型的参数, 直至训练样本集中的所有训练样本训练结束, 完成一次迭代;

设定全卷积神经网络模型的损失函数：

$$LOSS = -\frac{1}{n} \sum_k^n [o_k \ln y_k + (1 - o_k) \ln (1 - y_k)] \quad (5)$$

其中，n 表示输出层的节点数， o_k 表示输出层第 k 个节点的实际值， y_k 表示输出层第 k 个节点的输出值；

5 判断训练是否满足结束条件，如果满足结束条件，则结束训练，输出经过训练的全卷积神经网络模型，如果不满足结束条件，将继续训练模型，其中，所述结束条件包括第一结束条件或/和第二结束条件中的一个或两个，第一结束条件为当前迭代次数大于设定的最大迭代次数，第二结束条件为连续多次迭代时损失函数值的变化小于设定目标值。

10 为了实现上述目的，本申请的再一个方面是提供一种计算机可读存储介质，所述计算机可读存储介质中包括语音增强程序，所述语音增强程序被处理器执行时，实现如上所述的语音增强方法的步骤。

相对于现有技术，本申请具有以下优点和有益效果：

15 本申请通过构建全卷积神经网络模型作为语音增强模型，输入原始语音信号进行处理，得到增强语音信号。全卷积神经网络模型中，删除了全连接层，仅包含卷积层，大大减小了神经网络的参数，使得全卷积神经网络模型可以适用于限制内存的移动设备中，并且每个输出样本仅仅依赖相邻输入，可以通过相关的更少的权重值很好地保留语音信号的原始信息和空间排列信息。

20

附图说明

图1为本申请所述基于全卷积神经网络的语音增强方法的流程示意图；

图2为本申请中全卷积神经网络模型结构示意图；

图3为本申请中语音增强程序的模块示意图。

25 本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

30 下面将参考附图来描述本申请所述的实施例。本领域的普通技术人员可以认识到，在不偏离本申请的精神和范围的情况下，可以用各种不同的方式或其组合对所描述的实施例进行修正。因此，附图和描述在本质上是说明性

的，仅仅用以解释本申请，而不是用于限制权利要求的保护范围。此外，在本说明书中，附图未按比例画出，并且相同的附图标记表示相同的部分。

图 1 为本申请所述基于全卷积神经网络的语音增强方法的流程示意图，如图 1 所示，本申请所述基于全卷积神经网络的语音增强方法包括以下步骤：

5 步骤 S1、构建全卷积神经网络模型，如图 2 所示，所述全卷积神经网络模型包括输入层、隐含层和输出层，所述隐含层为多个卷积层，每个卷积层均具有多个滤波器，所述输出层的输出模型为：

$$y_t = F^T * R_t \quad (1)$$

10 其中， y_t 是输出层的第 t 个节点， F^T 是滤波器的权重矩阵的转置， $F \in R^{f \times 1}$ (f 表示滤波器尺寸)， R_t 是隐含层的第 t 个节点；

步骤 S2、训练所述全卷积神经网络模型；

步骤 S3、将原始语音信号输入经过训练的全卷积神经网络模型；

步骤 S4、输出增强语音信号。

15 本申请中，滤波器的权重矩阵 F 在卷积操作过程中是共享的，因此，无论输出层节点是高频部分还是低频部分，隐含层节点 R_t 与相邻的两个节点 R_{t-1} 和 R_{t+1} 不会很相似，隐含层节点与相邻节点是否相似取决于原始的输入层节点的输入，使得全卷积神经网络可以很好地保留原始输入信息。

20 本申请中通过构建全卷积神经网络模型作为语音增强模型，输入原始语音信号进行处理，得到增强语音信号。全卷积神经网络模型中，删除了全连接层，仅包含卷积层，大大减小了神经网络的参数，使得全卷积神经网络模型可以适应限制内存的移动设备中，例如手机等移动终端，并且每个输出样本仅仅依赖相邻输入，可以通过相关的更少的权重值很好地保留语音信号的原始信息和空间排列信息。

25 本申请的一个可选实施例中，所述全卷积神经网络模型包括：输入层、六个卷积层(具有 padding)和输出层，每个卷积层均具有 1024 个节点，卷积跨度为 1，每个卷积层均具有 15 个尺寸为 11 的滤波器，根据下式构建全卷积神经网络模型的隐含层的模型：

$$h_j^1 = f(\sum_{i=1}^n (w_{ik}^1 x_i) + a_k^1) \quad (2)$$

$$h_k^l = f(\sum_{j=1}^H (w_{kj}^l h_j^{l-1}) + b_k^l) \quad l \in (2, 3, \dots, L) \quad (3)$$

30 其中， h_j^1 表示第 1 个隐含层的第 j 个节点的输出值， x_i 表示输入层的第 i 个节点的变量， w_{ik}^1 表示输入层第 i 个节点和第 1 个隐含层第 k 个节点的连接

权重值, a_k^1 表示第 1 个隐含层第 k 个节点的偏移量, n 表示输入层的节点数, h_k^l 表示第 l 个隐含层的第 k 个节点的输出值, h_j^{l-1} 表示第 l-1 个隐含层的第 j 个节点的输出值, w_{kj}^l 表示第 l 个隐含层的第 k 个节点和第 l-1 个隐含层的第 j 个节点的连接权重值, b_k^l 表示第 l 个隐含层第 k 个节点的偏移量, H 为隐含层的节点数, f 为激励函数, 选择 PReLU 激活函数。

本申请的一个可选实施例中, 训练所述全卷积神经网络模型包括:

对所述全卷积神经网络模型的参数进行初始赋值, 所述参数包括输入层和隐含层的连接权重值、相邻隐含层之间的连接权重值和隐含层的偏移量;

构建样本集, 并将所述样本集按比例划分为训练样本集和测试样本集, 其中, 样本集中的样本可以从 TIMIT 语料库中随机选取, 训练样本集和测试样本集中样本个数的比例为 6: 1, 例如, 从 TIMIT 语料库中随机选取 700 个短语, 其中的 600 个短语构成训练样本集, 其余的 100 个短语构成测试样本集, 训练样本集中包含 5 种信噪比下的 5 种噪声类型 (白噪声、粉噪声、办公室噪声、超市噪声和街道噪声), 测试样本集中包含与训练样本集中相同或不同的信噪比和噪声类型, 信噪比可以不同, 且噪声类型也可以不同, 以使测试条件更加贴近真实。本申请中的训练样本集中仅列举出 5 种噪声类型, 但本申请并不限于此。

输入所述训练样本集中的一个训练样本, 并从所述训练样本中提取对数功率谱 (LPS, Log power spectra) 特征向量; 例如, 在输入的训练样本中, 选取原始语音的 512 个采样点作为一帧, 并且, 每帧提取 257 维 LPS 向量作为特征向量。

将训练样本的输入数据代入公式 (1) - (3), 计算隐含层各节点的输出值和输出层各节点的输出值;

计算输出层各节点误差:

$$e_k = o_k - y_k \quad (4)$$

其中, e_k 表示输出层第 k 个节点的误差, o_k 表示输出层第 k 个节点的实际值, y_k 表示输出层第 k 个节点的输出值;

基于误差反向传播更新所述全卷积神经网络模型的参数;

输入下一个训练样本, 继续更新全卷积神经网络模型的参数, 直至训练样本集中的所有训练样本训练结束, 完成一次迭代;

设定全卷积神经网络模型的损失函数:

$$LOSS = -\frac{1}{n} \sum_k^n [o_k \ln y_k + (1 - o_k) \ln (1 - y_k)] \quad (5)$$

其中，n 表示输出层的节点数， o_k 表示输出层第 k 个节点的实际值， y_k 表示输出层第 k 个节点的输出值；

判断训练是否满足结束条件，如果满足结束条件，则结束训练，输出经过训练的全卷积神经网络模型，如果不满足结束条件，将继续训练模型，其中，所述结束条件包括第一结束条件或/和第二结束条件中的一个或两个，第一结束条件为当前迭代次数大于设定的最大迭代次数，第二结束条件为连续多次迭代时损失函数值的变化小于设定目标值。

优选地，根据下式计算测试误差：

$$MSE = \frac{1}{2N} \sum_{Z=1}^N \sum_{k=1}^n (O'_k{}^Z - y'_k{}^Z)^2 \quad (6)$$

其中，MSE 表示测试误差，N 表示测试样本集的样本个数， $O'_k{}^Z$ 表示测试样本集的样本 z 在输出层第 k 个节点的实际值， $y'_k{}^Z$ 表示测试样本集的样本 z 在输出层第 k 个节点的输出值。测试误差越小，表明构建的全卷积神经网络模型的精度越高。

本申请中，对全卷积神经网络模型的输出数据进行归一化处理，之后，再进行输出层节点误差的计算和测试误差的计算等，以减小测试误差，提高模型精度。

优选地，通过语音质量评价（PESQ, Perceptual evaluation of speech quality）评价语音质量，通过短时客观可懂度得分（STOI, Short Time Objective Intelligibility）评价语音的可懂度。

通过本申请的全卷积神经网络模型进行语音增强，相对于包含全连接层的深度神经网络模型和卷积神经网络模型，PESQ 和 STOI 均有所提高，PESQ 可以提高 0.5 左右，STOI 可以提高 0.2-0.3 左右。

优选地，隐含层的模型应用 PReLU 激活函数。

优选地，训练所述全卷积神经网络模型采用 TIMIT 语料库，将其划分为训练集和测试集。

优选地，隐含层的模型应用 Adam 来最小化纯净语音和增强语音的最小均方误差。

优选地，输出增强语音信号通过 PESQ 和短时客观可懂度得分 STOI 来判断增强质量。

本申请所述基于全卷积神经网络的语音增强方法应用于电子装置，电子装置可以是电视机、智能手机、平板电脑、计算机等终端设备。然而，电子装置并不限于所列举示例，电子装置可以是用户控制的通过语音识别技术处理用户命令的任何其他装置，通过对输入用户的语音进行语音增强处理，输出语音识别结果。

所述电子装置包括：存储器和处理器，所述存储器中包括语音增强程序，所述语音增强程序被所述处理器执行时实现如下步骤：

构建全卷积神经网络模型，所述全卷积神经网络模型包括输入层、隐含层和输出层，所述隐含层为多个卷积层，每个卷积层均具有多个滤波器，所述输出层的输出模型为：

$$y_t = F^T * R_t \quad (1)$$

其中， y_t 是输出层的第 t 个节点， F^T 是滤波器的权重矩阵的转置， $F \in R^{f \times 1}$ ， f 表示滤波器尺寸， R_t 是隐含层的第 t 个节点；

训练所述全卷积神经网络模型；

将原始语音信号输入经过训练的全卷积神经网络模型；

输出增强语音信号。

存储器包括至少一种类型的可读存储介质，可以是闪存、硬盘、光盘等非易失性存储介质，也可以是插接式硬盘等，且并不限于此，可以是以非暂时性方式存储指令或软件以及任何相关联的数据文件并向处理器提供指令或软件程序以使该处理器能够执行指令或软件程序的任何装置。

所述电子装置还包括语音接收器，通过电子装置的麦克风等设备接收用户的语音信号，再对输入的语音信号进行语音增强处理。

处理器可以是中央处理器、微处理器或其他数据处理芯片等，可以运行存储器中的存储程序。

处理器通过梯度下降法寻找全卷积神经网络中各滤波器的权重值。

本申请的一个可选实施例中，所述全卷积神经网络模型中隐含层的模型为：

$$h_j^1 = f(\sum_{i=1}^n (w_{ik}^1 x_i) + a_k^1) \quad (2)$$

$$h_k^l = f(\sum_{j=1}^H (w_{kj}^l h_j^{l-1}) + b_k^l) \quad l \in (2, 3, \dots, L) \quad (3)$$

其中， h_j^1 表示第 1 个隐含层的第 j 个节点的输出值， x_i 表示输入层的第 i 个节点的变量， w_{ik}^1 表示输入层第 i 个节点和第 1 个隐含层第 k 个节点的连接

权重值, a_k^1 表示第 1 个隐含层第 k 个节点的偏移量, n 表示输入层的节点数, h_k^l 表示第 l 个隐含层的第 k 个节点的输出值, h_j^{l-1} 表示第 l-1 个隐含层的第 j 个节点的输出值, w_{kj}^l 表示第 l 个隐含层的第 k 个节点和第 l-1 个隐含层的第 j 个节点的连接权重值, b_k^l 表示第 l 个隐含层第 k 个节点的偏移量, H 为隐含层的节点数, f 为激励函数, 其中, 激励函数可以选择 PReLU 激活函数、Sigmoid 函数、tanh 函数、Relu 函数等函数。

本申请的一个实施例中, 处理器训练所述全卷积神经网络模型的步骤包括:

对所述全卷积神经网络模型的参数进行初始赋值, 所述参数包括输入层和隐含层的连接权重值、相邻隐含层之间的连接权重值和隐含层的偏移量;

构建样本集, 并将所述样本集按比例划分为训练样本集和测试样本集, 其中, 样本集中的样本可以从 TIMIT 语料库中随机选取, 训练样本集和测试样本集中样本个数的比例为 6: 1, 例如, 从 TIMIT 语料库中随机选取 700 个短语, 其中的 600 个短语构成训练样本集, 其余的 100 个短语构成测试样本集, 训练样本集中包含 5 种信噪比下的 5 种噪声类型 (白噪声、粉噪声、办公室噪声、超市噪声和街道噪声), 测试样本集中包含与训练样本集中相同或不同的信噪比和噪声类型, 以使测试条件更加贴近真实。本申请中的训练样本集中仅列举出 5 种噪声类型, 但本申请并不限于此;

输入所述训练样本集中的一个训练样本, 并从所述训练样本中提取特征向量;

将训练样本的输入数据代入公式 (1) - (3), 计算隐含层各节点的输出值和输出层各节点的输出值;

计算输出层各节点误差:

$$e_k = o_k - y_k \quad (4)$$

其中, e_k 表示输出层第 k 个节点的误差, o_k 表示输出层第 k 个节点的实际值, y_k 表示输出层第 k 个节点的输出值;

基于误差反向传播更新所述全卷积神经网络模型的参数;

输入下一个训练样本, 继续更新全卷积神经网络模型的参数, 直至训练样本集中的所有训练样本训练结束, 完成一次迭代;

设定全卷积神经网络模型的损失函数:

$$LOSS = -\frac{1}{n} \sum_k^n [o_k \ln y_k + (1 - o_k) \ln (1 - y_k)] \quad (5)$$

其中, n 表示输出层的节点数, o_k 表示输出层第 k 个节点的实际值, y_k 表示输出层第 k 个节点的输出值;

判断训练是否满足结束条件, 如果满足结束条件, 则结束训练, 输出经过训练的全卷积神经网络模型, 如果不满足结束条件, 将继续训练模型, 其中, 所述结束条件包括第一结束条件或/和第二结束条件中的一个或两个, 第一结束条件为当前迭代次数大于设定的最大迭代次数, 第二结束条件为连续多次迭代时损失函数值的变化小于设定目标值。

优选地, 根据下式计算测试误差:

$$MSE = \frac{1}{2N} \sum_{Z=1}^N \sum_{k=1}^n (O_k^Z - y_k^Z)^2 \quad (6)$$

其中, MSE 表示测试误差, N 表示测试样本集的样本个数, O_k^Z 表示测试样本集的样本 z 在输出层第 k 个节点的实际值, y_k^Z 表示测试样本集的样本 z 在输出层第 k 个节点的输出值。

在其他实施例中, 语音增强程序还可以被分割为一个或者多个模块, 一个或者多个模块被存储于存储器中, 并由处理器执行, 以完成本申请。本申请所称的模块是指能够完成特定功能的一系列计算机程序指令段。所述语音增强程序可以被分割为: 模型构建模块 1、模型训练模块 2、输入模块 3 和输出模块 4。上述模块所实现的功能或操作步骤均与上文类似, 此处不再详述, 示例性地, 例如其中:

模型构建模块 1, 构建全卷积神经网络模型, 所述全卷积神经网络模型包括输入层、隐含层和输出层, 所述隐含层为多个卷积层, 每个卷积层均具有多个滤波器, 所述输出层的输出模型为:

$$y_t = F^T * R_t \quad (1)$$

其中, t 是节点的索引, y_t 是输出层的第 t 个节点, F 是滤波器, $F \in R^{f \times 1}$, f 表示滤波器尺寸, R_t 是隐含层的第 t 个节点;

模型训练模块 2, 训练所述全卷积神经网络模型;

输入模块 3, 将原始语音信号输入经过训练的全卷积神经网络模型;

输出模块 4, 输出增强语音信号。

优选地, 隐含层的模型应用 PReLU 激活函数。

优选地, 训练所述全卷积神经网络模型采用 TIMIT 语料库, 将其划分为训练集和测试集。

优选地，隐含层的模型应用 Adam 来最小化纯净语音和增强语音的最小均方误差。

优选地，输出增强语音信号通过 PESQ 和短时客观可懂度得分 STOI 来判断增强质量。

5 本申请的一个实施例中，计算机可读存储介质可以是任何包含或存储程序或指令的有形介质，其中的程序可以被执行，通过存储的程序指令相关的硬件实现相应的功能。例如，计算机可读存储介质可以是计算机磁盘、硬盘、随机存取存储器、只读存储器等。本申请并不限于此，可以是以非暂时性方式存储指令或软件以及任何相关数据文件或数据结构并且可提供给处理器以
10 使处理器执行其中的程序或指令的任何装置。所述计算机可读存储介质中包括语音增强程序，所述语音增强程序被处理器执行时，实现如下的语音增强方法：

构建全卷积神经网络模型，所述全卷积神经网络模型包括输入层、隐含层和输出层，所述隐含层为多个卷积层，每个卷积层均具有多个滤波器，所述输出层的输出模型为：
15

$$y_t = F^T * R_t \quad (1)$$

其中， y_t 是输出层的第 t 个节点， F^T 是滤波器的权重矩阵的转置， $F \in R^{f \times 1}$ ， f 表示滤波器尺寸， R_t 是隐含层的第 t 个节点；

训练所述全卷积神经网络模型；

20 将原始语音信号输入经过训练的全卷积神经网络模型；
输出增强语音信号。

优选地，根据下式构建全卷积神经网络模型的隐含层的模型：

$$h_j^1 = f(\sum_{i=1}^n (w_{ik}^1 x_i) + a_k^1) \quad (2)$$

$$h_k^l = f(\sum_{j=1}^H (w_{kj}^l h_j^{l-1}) + b_k^l) \quad l \in (2, 3, \dots, L) \quad (3)$$

25 其中， h_j^1 表示第 1 个隐含层的第 j 个节点的输出值， x_i 表示输入层的第 i 个节点的变量， w_{ik}^1 表示输入层第 i 个节点和第 1 个隐含层第 k 个节点的连接权重值， a_k^1 表示第 1 个隐含层第 k 个节点的偏移量， n 表示输入层的节点数， h_k^l 表示第 1 个隐含层的第 k 个节点的输出值， h_j^{l-1} 表示第 $l-1$ 个隐含层的第 j 个节点的输出值， w_{kj}^l 表示第 1 个隐含层的第 k 个节点和第 $l-1$ 个隐含层的第 j 个节点的连接权重值， b_k^l 表示第 1 个隐含层第 k 个节点的偏移量， H 为隐含层的节点数， f 为激励函数。
30

优选地，训练所述全卷积神经网络模型包括：

对所述全卷积神经网络模型的参数进行初始赋值，所述参数包括输入层和隐含层的连接权重值、相邻隐含层之间的连接权重值和隐含层的偏移量；

构建样本集，并将所述样本集按比例划分为训练样本集和测试样本集；

5 输入所述训练样本集中的一个训练样本，并从所述训练样本中提取特征向量；

将训练样本的输入数据代入公式（1）-（3），计算隐含层各节点的输出值和输出层各节点的输出值；

计算输出层各节点误差：

$$10 \quad e_k = o_k - y_k \quad (4)$$

其中， e_k 表示输出层第 k 个节点的误差， o_k 表示输出层第 k 个节点的实际值， y_k 表示输出层第 k 个节点的输出值；

基于误差反向传播更新所述全卷积神经网络模型的参数；

15 输入下一个训练样本，继续更新全卷积神经网络模型的参数，直至训练样本集中的所有训练样本训练结束，完成一次迭代；

设定全卷积神经网络模型的损失函数：

$$LOSS = -\frac{1}{n} \sum_k^n [o_k \ln y_k + (1 - o_k) \ln (1 - y_k)] \quad (5)$$

其中， n 表示输出层的节点数， o_k 表示输出层第 k 个节点的实际值， y_k 表示输出层第 k 个节点的输出值；

20 判断训练是否满足结束条件，如果满足结束条件，则结束训练，输出经过训练的全卷积神经网络模型，如果不满足结束条件，将继续训练模型，其中，所述结束条件包括第一结束条件或/和第二结束条件中的一个或两个，第一结束条件为当前迭代次数大于设定的最大迭代次数，第二结束条件为连续多次迭代时损失函数值的变化小于设定目标值。

25 优选地，根据下式计算测试误差：

$$MSE = \frac{1}{2N} \sum_{Z=1}^N \sum_{k=1}^n (O'_k{}^Z - y'_k{}^Z)^2 \quad (6)$$

其中， MSE 表示测试误差， N 表示测试样本集的样本个数， $O'_k{}^Z$ 表示测试样本集的样本 z 在输出层第 k 个节点的实际值， $y'_k{}^Z$ 表示测试样本集的样本 z 在输出层第 k 个节点的输出值。

30 优选地，测试样本集中的测试样本与训练样本集中的训练样本的信噪比

和噪声类型不同。

优选地，所述全卷积神经网络模型包括输入层、六个卷积层和输出层，每个卷积层均具有 1024 个节点，卷积跨度为 1。

5 本申请之计算机可读存储介质的具体实施方式与上述语音增强方法、电子装置的具体实施方式大致相同，在此不再赘述。

需要说明的是，在本文中，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、装置、物品或者方法不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、装置、物品或者方法所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括该要素的过程、装置、物品或者方法中还存在另外的相同要素。

上述本申请实施例序号仅仅为了描述，不代表实施例的优劣。通过以上的实施方式的描述，本领域的技术人员可以清楚地了解到上述实施例方法可借助软件加必需的通用硬件平台的方式来实现，当然也可以通过硬件，但很多情况下前者是更佳的实施方式。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来，该计算机软件产品存储在如上所述的一个存储介质(如 ROM/RAM、磁碟、光盘)中，包括若干指令用以使得一台终端设备(可以是手机，计算机，服务器，或者网络设备等)执行本申请各个实施例所述的方法。

20 以上仅为本申请的优选实施例，并非因此限制本申请的专利范围，凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其他相关的技术领域，均同理包括在本申请的专利保护范围内。

权 利 要 求 书

1. 一种基于全卷积神经网络的语音增强方法，应用于电子装置，其特征在于，

构建全卷积神经网络模型，所述全卷积神经网络模型包括输入层、隐含层和输出层，所述隐含层为多个卷积层，每个卷积层均具有多个滤波器，所述输出层的输出模型为：

$$y_t = F^T * R_t \quad (1)$$

其中， y_t 是输出层的第 t 个节点， F^T 是滤波器的权重矩阵的转置， $F \in R^{f \times 1}$ ， f 表示滤波器尺寸， R_t 是隐含层的第 t 个节点；

训练所述全卷积神经网络模型；

将原始语音信号输入经过训练的全卷积神经网络模型；

输出增强语音信号。

2. 根据权利要求 1 所述的基于全卷积神经网络的语音增强方法，其特征在于，根据下式构建全卷积神经网络模型的隐含层的模型：

$$h_j^1 = f(\sum_{i=1}^n (w_{ik}^1 x_i) + a_k^1) \quad (2)$$

$$h_k^l = f(\sum_{j=1}^H (w_{kj}^l h_j^{l-1}) + b_k^l) \quad l \in (2, 3, \dots, L) \quad (3)$$

其中， h_j^1 表示第 1 个隐含层的第 j 个节点的输出值， x_i 表示输入层的第 i 个节点的变量， w_{ik}^1 表示输入层第 i 个节点和第 1 个隐含层第 k 个节点的连接权重值， a_k^1 表示第 1 个隐含层第 k 个节点的偏移量， n 表示输入层的节点数， h_k^l 表示第 l 个隐含层的第 k 个节点的输出值， h_j^{l-1} 表示第 $l-1$ 个隐含层的第 j 个节点的输出值， w_{kj}^l 表示第 l 个隐含层的第 k 个节点和第 $l-1$ 个隐含层的第 j 个节点的连接权重值， b_k^l 表示第 l 个隐含层第 k 个节点的偏移量， H 为隐含层的节点数， f 为激励函数。

3. 根据权利要求 2 所述的基于全卷积神经网络的语音增强方法，其特征在于，训练所述全卷积神经网络模型包括：

对所述全卷积神经网络模型的参数进行初始赋值，所述参数包括输入层和隐含层的连接权重值、相邻隐含层之间的连接权重值和隐含层的偏移量；

构建样本集，并将所述样本集按比例划分为训练样本集和测试样本集；

输入所述训练样本集中的一个训练样本，并从所述训练样本中提取特征

向量；

将训练样本的输入数据代入公式 (1) - (3)，计算隐含层各节点的输出

值和输出层各节点的输出值；

计算输出层各节点误差：

$$e_k = o_k - y_k \quad (4)$$

其中， e_k 表示输出层第 k 个节点的误差， o_k 表示输出层第 k 个节点的实际值， y_k 表示输出层第 k 个节点的输出值；

基于误差反向传播更新所述全卷积神经网络模型的参数；

输入下一个训练样本，继续更新全卷积神经网络模型的参数，直至训练样本集中的所有训练样本训练结束，完成一次迭代；

设定全卷积神经网络模型的损失函数：

$$LOSS = -\frac{1}{n} \sum_k^n [o_k \ln y_k + (1 - o_k) \ln (1 - y_k)] \quad (5)$$

其中， n 表示输出层的节点数， o_k 表示输出层第 k 个节点的实际值， y_k 表示输出层第 k 个节点的输出值；

判断训练是否满足结束条件，如果满足结束条件，则结束训练，输出经过训练的全卷积神经网络模型，如果不满足结束条件，将继续训练模型，其中，所述结束条件包括第一结束条件或/和第二结束条件中的一个或两个，第一结束条件为当前迭代次数大于设定的最大迭代次数，第二结束条件为连续多次迭代时损失函数值的变化小于设定目标值。

4. 根据权利要求 3 所述的基于全卷积神经网络的语音增强方法，其特征在于，根据下式计算测试误差：

$$MSE = \frac{1}{2N} \sum_{z=1}^N \sum_{k=1}^n (O_k^z - y_k^z)^2 \quad (6)$$

其中， MSE 表示测试误差， N 表示测试样本集的样本个数， O_k^z 表示测试样本集的样本 z 在输出层第 k 个节点的实际值， y_k^z 表示测试样本集的样本 z 在输出层第 k 个节点的输出值。

5. 根据权利要求 3 所述的基于全卷积神经网络的语音增强方法，其特征在于，测试样本集中的测试样本与训练样本集中的训练样本的信噪比不同，且噪声类型也不同。

6. 根据权利要求 3 所述的基于全卷积神经网络的语音增强方法，其特征在于，通过梯度下降法寻找全卷积神经网络中各滤波器的权重值。

7. 根据权利要求 3 所述的基于全卷积神经网络的语音增强方法，其特征在于，所述隐含层的模型应用 PReLU 激活函数。

8. 根据权利要求 2 所述的基于全卷积神经网络的语音增强方法,其特征
在于,训练所述全卷积神经网络模型采用 TIMIT 语料库,将其划分为训练集
和测试集。

9. 根据权利要求 3 所述的基于全卷积神经网络的语音增强方法,其特征
5 在于,所述隐含层的模型应用 Adam 来最小化纯净语音和增强语音的最小均方
误差。

10. 根据权利要求 3 所述的基于全卷积神经网络的语音增强方法,其特
征在于,输出增强语音信号通过 PESQ 和短时客观可懂度得分 STOI 来判断增
强质量。

11. 根据权利要求 1 至 10 中任一项所述的基于全卷积神经网络的语音增
10 强方法,其特征在于,所述全卷积神经网络模型包括输入层、六个卷积层和
输出层,每个卷积层均具有 1024 个节点,卷积跨度为 1。

12. 一种电子装置,其特征不在于,该电子装置包括:存储器和处理器,
所述存储器中包括语音增强程序,所述语音增强程序被所述处理器执行时实
15 现如下步骤:

构建全卷积神经网络模型,所述全卷积神经网络模型包括输入层、隐含
层和输出层,所述隐含层为多个卷积层,每个卷积层均具有多个滤波器,所
述输出层的输出模型为:

$$y_t = F^T * R_t \quad (1)$$

20 其中, y_t 是输出层的第 t 个节点, F^T 是滤波器的权重矩阵的转置, $F \in R^{f \times 1}$, f 表示滤波器尺寸, R_t 是隐含层的第 t 个节点;

训练所述全卷积神经网络模型;

将原始语音信号输入经过训练的全卷积神经网络模型;

输出增强语音信号。

25 13. 根据权利要求 12 所述的电子装置,其特征不在于,所述全卷积神经网络
模型中隐含层的模型为:

$$h_j^1 = f(\sum_{i=1}^n (w_{ik}^1 x_i) + a_k^1) \quad (2)$$

$$h_k^l = f(\sum_{j=1}^H (w_{kj}^l h_j^{l-1}) + b_k^l) \quad l \in (2, 3, \dots, L) \quad (3)$$

30 其中, h_j^1 表示第 1 个隐含层的第 j 个节点的输出值, x_i 表示输入层的第
 i 个节点的变量, w_{ik}^1 表示输入层第 i 个节点和第 1 个隐含层第 k 个节点的连
接权重值, a_k^1 表示第 1 个隐含层第 k 个节点的偏移量, n 表示输入层的节点

数, h_k^l 表示第 1 个隐含层的第 k 个节点的输出值, h_j^{l-1} 表示第 1-1 个隐含层的第 j 个节点的输出值, w_{kj}^l 表示第 1 个隐含层的第 k 个节点和第 1-1 个隐含层的第 j 个节点的连接权重值, b_k^l 表示第 1 个隐含层第 k 个节点的偏移量, H 为隐含层的节点数, f 为激励函数。

5 14. 根据权利要求 12 所述的电子装置, 其特征在于, 处理器训练所述全卷积神经网络模型包括:

对所述全卷积神经网络模型的参数进行初始赋值, 所述参数包括输入层和隐含层的连接权重值、相邻隐含层之间的连接权重值和隐含层的偏移量;

构建样本集, 并将所述样本集按比例划分为训练样本集和测试样本集;

10 输入所述训练样本集中的一个训练样本, 并从所述训练样本中提取特征向量;

将训练样本的输入数据代入公式 (1) - (3), 计算隐含层各节点的输出值和输出层各节点的输出值;

计算输出层各节点误差:

15
$$e_k = o_k - y_k \quad (4)$$

其中, e_k 表示输出层第 k 个节点的误差, o_k 表示输出层第 k 个节点的实际值, y_k 表示输出层第 k 个节点的输出值;

基于误差反向传播更新所述全卷积神经网络模型的参数;

20 输入下一个训练样本, 继续更新全卷积神经网络模型的参数, 直至训练样本集中的所有训练样本训练结束, 完成一次迭代;

设定全卷积神经网络模型的损失函数:

$$LOSS = -\frac{1}{n} \sum_k^n [o_k \ln y_k + (1 - o_k) \ln (1 - y_k)] \quad (5)$$

其中, n 表示输出层的节点数, o_k 表示输出层第 k 个节点的实际值, y_k 表示输出层第 k 个节点的输出值;

25 判断训练是否满足结束条件, 如果满足结束条件, 则结束训练, 输出经过训练的全卷积神经网络模型, 如果不满足结束条件, 将继续训练模型, 其中, 所述结束条件包括第一结束条件或/和第二结束条件中的一个或两个, 第一结束条件为当前迭代次数大于设定的最大迭代次数, 第二结束条件为连续多次迭代时损失函数值的变化小于设定目标值。

30 15. 根据权利要求 12 所述的电子装置, 其特征在于, 处理器通过梯度下降法寻找全卷积神经网络中各滤波器的权重值。

16. 根据权利要求 12 所述的电子装置, 其特征在于, 所述隐含层的模型应用 PReLU 激活函数。

17. 根据权利要求 12 所述的电子装置, 其特征在于, 训练所述全卷积神经网络模型采用 TIMIT 语料库, 将其划分为训练集和测试集。

5 18. 根据权利要求 12 所述的电子装置, 其特征在于, 所述隐含层的模型应用 Adam 来最小化纯净语音和增强语音的最小均方误差。

19. 根据权利要求 12 所述的电子装置, 其特征在于, 输出增强语音信号后通过 PESQ 和短时客观可懂度得分 STOI 来判断语音增强质量。

10 20. 一种计算机可读存储介质, 其特征在于, 所述计算机可读存储介质中包括语音增强程序, 所述语音增强程序被处理器执行时, 实现如权利要求 1 至 10 中任一项所述的语音增强方法的步骤。

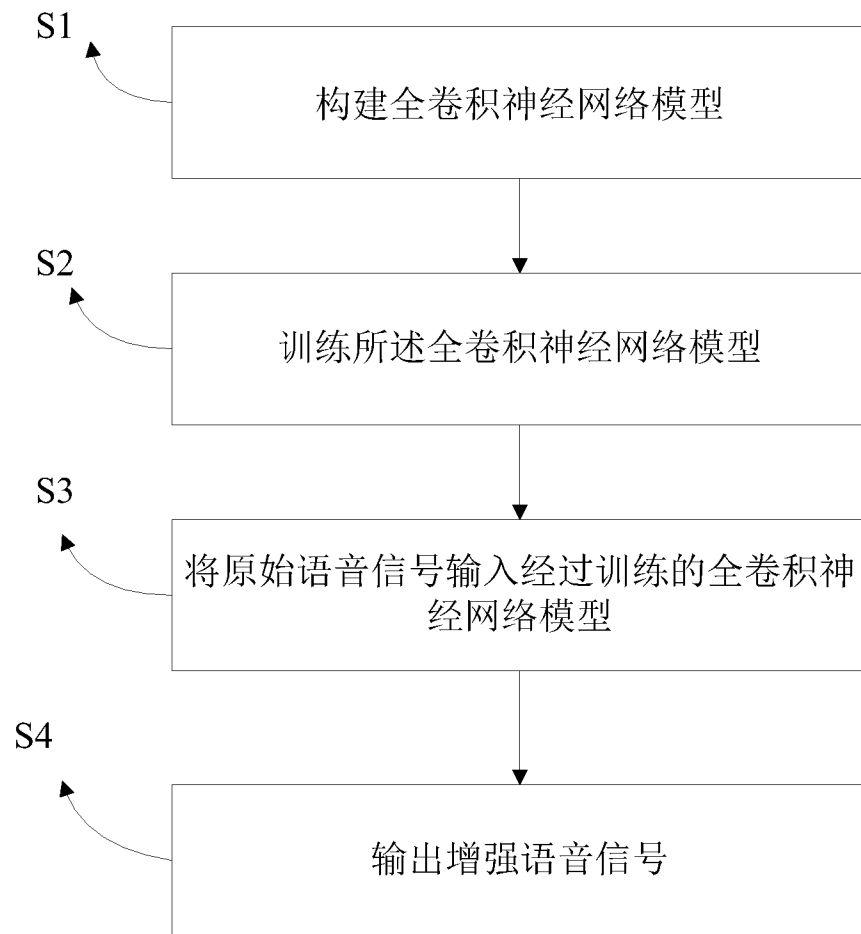


图 1

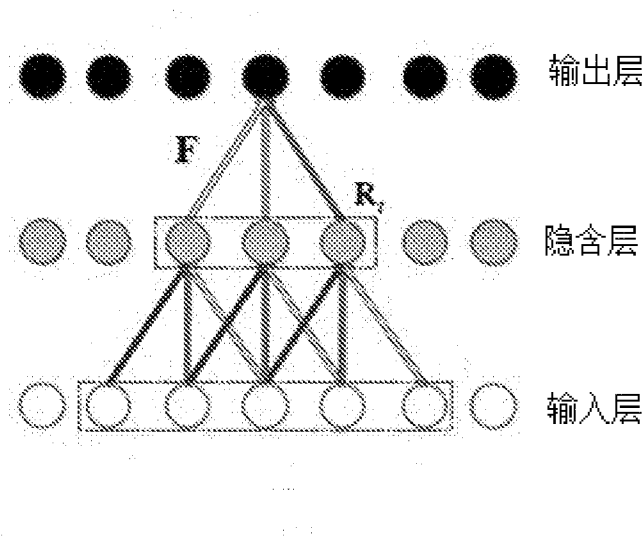


图 2

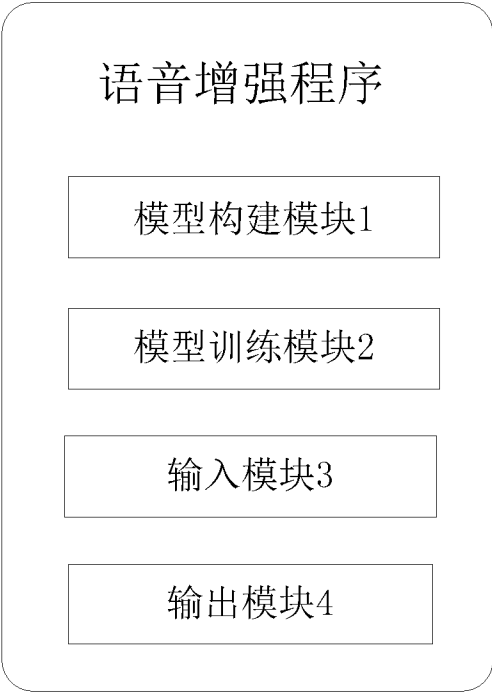


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/089180

A. CLASSIFICATION OF SUBJECT MATTER

G10L 21/02(2013.01)i; G10L 21/0208(2013.01)i; G10L 25/00(2013.01)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)

G10L

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)

CNABS; CNTXT; CNKI; 万方, WANFANG; 超星, CHAOXING; VEN; USTXT; EPTXT; IEEE: 语音, 增强, 隐含, 卷积, 输出层, 连接层, 卷积神经网络, 转置, 去掉, 删除, 节点, 权重, 矩阵, 去除, 隐藏层, voice, neural, network, convolution, output, matrix, weight

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 109326299 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 12 February 2019 (2019-02-12) description, paragraphs [0068]-[0170], and figures 1-3	1-20
A	CN 108172238 A (GUANGZHOU VOIBOOK TECHNOLOGY CO., LTD.) 15 June 2018 (2018-06-15) description, paragraphs [0018]-[0036], and figures 1-6	1-20
A	CN 106847302 A (DALIAN UNIVERSITY OF TECHNOLOGY) 13 June 2017 (2017-06-13) entire document	1-20
A	CN 107845389 A (BEIJING UNIVERSITY OF TECHNOLOGY) 27 March 2018 (2018-03-27) entire document	1-20
A	US 2016322055 A1 (GOOGLE INC.) 03 November 2016 (2016-11-03) entire document	1-20



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

“D” document cited by the applicant in the international application

“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

23 July 2019

Date of mailing of the international search report

21 August 2019

Name and mailing address of the ISA/CN

China National Intellectual Property Administration
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/089180

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	109326299	A	12 February 2019	None	
CN	108172238	A	15 June 2018	None	
CN	106847302	A	13 June 2017	None	
CN	107845389	A	27 March 2018	None	
US	2016322055	A1	03 November 2016	US 9697826 B2	04 July 2017

国际检索报告

国际申请号

PCT/CN2019/089180

A. 主题的分类

G10L 21/02(2013.01)i; G10L 21/0208(2013.01)i; G10L 25/00(2013.01)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)

G10L

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))

CNABS;CNTXT;CNKI;万方;超星;VEN;USTXT;EPTXT;IEEE:语音, 增强, 隐含, 卷积, 输出层, 连接层, 卷积神经网络, 转置, 去掉, 删除, 节点, 权重, 矩阵, 去除, 隐藏层, voice, neural, network, convolution, output, matrix, weight

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 109326299 A (平安科技深圳有限公司) 2019年 2月 12日 (2019 - 02 - 12) 说明书第[0068]-[0170]段, 附图1-3	1-20
A	CN 108172238 A (广州音书科技有限公司) 2018年 6月 15日 (2018 - 06 - 15) 说明书第[0018]-[0036]段, 附图1-6	1-20
A	CN 106847302 A (大连理工大学) 2017年 6月 13日 (2017 - 06 - 13) 全文	1-20
A	CN 107845389 A (北京工业大学) 2018年 3月 27日 (2018 - 03 - 27) 全文	1-20
A	US 2016322055 A1 (GOOGLE INC) 2016年 11月 3日 (2016 - 11 - 03) 全文	1-20

☐ 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, 但为了理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

2019年 7月 23日

国际检索报告邮寄日期

2019年 8月 21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

周群

电话号码 86-(0512)-88997268

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/089180

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	109326299	A	2019年 2月 12日	无	
CN	108172238	A	2018年 6月 15日	无	
CN	106847302	A	2017年 6月 13日	无	
CN	107845389	A	2018年 3月 27日	无	
US	2016322055	A1	2016年 11月 3日	US 9697826 B2	2017年 7月 4日