



(12)发明专利申请

(10)申请公布号 CN 109299142 A

(43)申请公布日 2019.02.01

(21)申请号 201811353444.8

(22)申请日 2018.11.14

(71)申请人 中山大学

地址 510275 广东省广州市海珠区新港西路135号

(72)发明人 卢宇彤 瞿毅力 郑馥丹 陈志广

(74)专利代理机构 湖南兆弘专利事务所(普通合伙) 43008

代理人 谭武艺

(51)Int.Cl.

G06F 16/2453(2019.01)

G06N 3/00(2006.01)

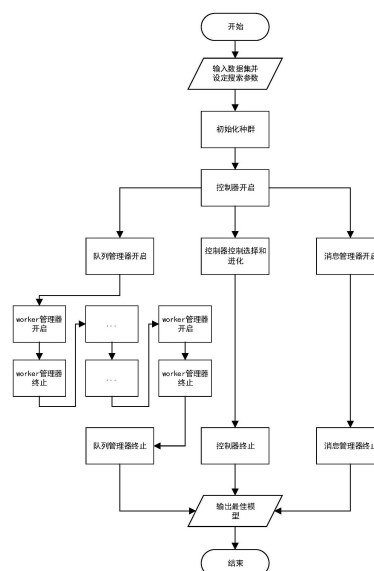
权利要求书3页 说明书10页 附图6页

(54)发明名称

一种基于进化算法的卷积神经网络结构搜索方法及系统

(57)摘要

本发明公开了一种基于进化算法的卷积神经网络结构搜索方法及系统,方法步骤包括输入数据集并设定预设参数,获取初始种群;通过作为主线程的控制器 T_c 将初始种群弹入队列 Q 并开启队列管理器 T_q 和消息管理器 T_m ,队列管理器 T_q 开启后针对队列 Q 中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_w 对其训练计算适应度,通过控制器 T_c 、队列管理器 T_q 、worker管理器 T_w 、消息管理器 T_m 的协同完成基于进化算法的卷积神经网络结构的并行搜索并输出最佳模型。本发明能针对给定数据集实现自动建模、调参和训练,具有性能高、规模化、流程化、确定性好的优点,尤其适用在高性能计算机集群上部署实施的优点。



1. 一种基于进化算法的卷积神经网络结构搜索方法,其特征在于实施步骤包括:

1) 输入数据集并设定预设参数;

2) 根据预设参数初始化种群,得到初始种群;

3) 通过作为主线程的控制器T_c将初始种群弹入队列Q并开启队列管理器T_q和消息管理器T_m,队列管理器T_q开启后针对队列Q中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器T_w对其训练计算适应度,且队列管理器T_q在完成训练的个体的个数达到预设参数中的总搜索模型个数后终止,且根据worker管理器T_w的状态更新字典W,字典W中记录全部worker的状态且状态为正在训练不可用或空闲可用;消息管理器T_m开启后如果收到worker管理器T_w发来的训练结果则将其更新字典M,且消息管理器T_m在收到训练结果达到预设参数中的总搜索模型个数后终止;

4) 控制器T_c监控字典M、队列Q和字典W;当字典M中有足够的可选个体、队列Q为空、字典W中有指定数量的可用worker,则通过worker进行一次优秀个体选择,选择策略为预设参数中的竞标赛选择策略;然后根据预设参数中的渐进式搜索设置,确定是否要向选择出的优秀个体中添加基因,再根据预设参数中的交配比例、变异概率、染色体分离进化还是组合进化以及基因冲突时的重进化策略,对选择出来的高适应度个体执行进化,进化之后的个体存入队列Q中;当控制器T_c监控到字典M中个体的个数达到预设参数中的总搜索模型个数时,跳转执行步骤5);

5) 输出最佳模型,结束。

2. 根据权利要求1所述的基于进化算法的卷积神经网络结构搜索方法,其特征在于,步骤2)的详细步骤包括:

2.1) 检测是否有人工设置的优秀染色体,如果有则读入,再根据预先设置的初始种群大小随机生成额外需要的染色体个数;

2.2) 先从可选算子列表O中随机选择出设定数量的算子组成算子染色体,再根据预设的策略生成连接基因,再组成染色体,连接染色体生成后与算子染色体通过一个染色体融合函数即可融合为完整的模型染色体,得到初始种群;

2.3) 针对初始种群中每一个染色体进行染色体合理性检测,如果未通过则需要舍弃重新随机生成新的染色体,最终得到通过染色体合理性检测的初始种群。

3. 根据权利要求1所述的基于进化算法的卷积神经网络结构搜索方法,其特征在于,步骤3)中队列管理器T_q开启后针对队列Q中的未训练染色体将其弹出解码的详细步骤包括:首先将染色体中的算子基因解码成预设的可选算子列表O中的算子名称,再根据连接基因将这些算子名称存入一个多层字典,最后再实现重复堆叠的组装过程,得到完整模型对应的一个字典;所述可选算子列表O存储已经被实现了的可以直接调用的每一个算子的算子名称,每个算子的类型和参数已经限定,且同一类型的两个算子如果有任何的参数不同,那么在可选算子列表O中两者会作为两个不同的算子存放;可选算子列表O确定且其中的算子均实现可解码后,每次搜索时每个个体中的全部算子均为可选算子列表O中的算子,同时每个个体中的算子的总个数也被限定;然后,根据预先设定的第三方模型训练工具遍历存储完整模型的字典,依次调用其中算子针对设定的训练工具的解码实现,得到在训练工具中此算子的字符串描述,根据模型字典将所有算子的字符串描述进行拼接,得到这个工具中模型的描述代码或者描述配置,再将其存储为训练工具可以读入的文件作为worker管理器

T_w 的输入。

4. 根据权利要求1所述的基于进化算法的卷积神经网络结构搜索方法,其特征在于,步骤3)中worker管理器 T_w 对其训练计算适应度的详细步骤包括:

3.1) 检查字典 W 中是否有空闲可用状态的worker,当且仅当字典 W 中有空闲可用状态的worker时跳转执行下一步;否则,继续检查字典 W ;

3.2) 选择worker,更新字典 W 将选中worker的状态设置为训练不可用;

3.3) 将解码后得到的模型提交到选中worker进行训练,并检测是否训练结束,当且仅当选worker训练结束时跳转执行下一步;

3.4) 更新字典 W 将选中worker的状态设置为空闲可用,释放选中worker;

3.5) 解析训练结果,调用如式(1)所示的适应度函数计算得到适应度;

$$f = w_a * a + w_m * m + w_t * t \quad (1)$$

式(1)中, f 为适应度, w_a 为训练精度 a 的权重函数, w_m 为模型参数量 m 的权重函数, w_t 为训练用时 t 的权重函数,其中训练精度 a 的权重函数 w_a 的值域为非负,模型参数量 m 的权重函数 w_m 的值域为0或负值,训练用时 t 的权重函数 w_t 的值域为0或负值;

3.6) 将适应度发送给消息管理器 T_m 。

5. 根据权利要求1所述的基于进化算法的卷积神经网络结构搜索方法,其特征在于,步骤4)中通过worker进行一次优秀个体选择时,选择的策略为下述方法之一:①、固定数量策略:每次选择可选群体中固定数目的优秀个体进行进化;②、固定比例策略:每次选择可选群体中固定比例的优秀个体进行进化;③、根据可用worker数确定,有多少worker可用则从可选群体中选择多少个体进化;步骤4)中对选择出来的高适应度个体执行进化时采用的进化操作为双点交叉、多点变异、重进化、分离进化与组合进化中的一种或者多种;其中,双点交叉是指从被选择的个体中随机挑选偶数个个体进行双点交叉操作,挑选比例即为交叉比例,将被挑选出的偶数个个体按每组两个个体分组,对每组两个个体的染色体执行双点交叉操作时,随机选择两个基因位点,两个染色体交换这两个位点间的基因片段;多点变异是指针对每个染色体进行多次变异,且每一次的变异概率相同或者不同,每一次变异时针对每个染色体,随机生成一个概率值作为变异概率,若大于预先设定的变异概率,则再随机选择一个位点对其进行变异,变异指0变为1或1变为0,这称为一次变异;重进化是指未通过染色体合理性检测的染色体可以根据预先设定进行再次进化或者重新进化;分离进化是指根据预先设定的参数选择直接对整条模型染色体进行进化,组合进化是指根据预先设定的参数选择将模型染色体分离成算子染色体和连接染色体分别进化。

6. 根据权利要求1所述的基于进化算法的卷积神经网络结构搜索方法,其特征在于,步骤3)中队列管理器 T_q 开启后针对队列 Q 中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_w 对其训练计算适应度后还包括进行早停策略控制步骤,所述早停策略控制的详细步骤包括:首先判断完成训练的个体的个数是否大于预设的早停迭代次数,如果完成训练的个体的个数大于预设的早停迭代次数、且训练结果中的适应度低于预设的适应度阈值,则队列管理器 T_q 提前终止、而不需要在完成训练的个体的个数达到预设参数中的总搜索模型个数后终止。

7. 根据权利要求1所述的基于进化算法的卷积神经网络结构搜索方法,其特征在于,步骤3)中队列管理器 T_q 开启后针对队列 Q 中的未训练染色体将其弹出解码后开启一个作为独

立临时线程的worker管理器Tw对其训练计算适应度时,worker训练计算适应度采用渐进式搜索策略,所述渐进式搜索策略的详细步骤包括:设定多级搜索数阈值,每个阈值表示已经开启训练的模型的个数,其中较低的阈值对应较小的搜索空间,较大的阈值对应较大的搜索空间,这里搜索空间大小的变化通过染色体中算子基因的个数来控制,使得初始种群在第一级搜索空间中构建并依次训练,之后直到已经开启训练的个体的个数达到了第一级阈值,此时对后续被开启训练的个体的染色体中的算子的个数进行统计,如果不满足第二级搜索空间的数量要求则对其进行算子添加操作,如此继续下去,直到在最后一级的搜索空间完成搜索;其中,进行算子添加操作的详细步骤包括:a)首先分离出算子染色体和连接染色体;b)然后往算子染色体的一个随机位点随机添加一个算子的基因;c)再在相同位点添加连接基因,其中的连接的有无也是随机的;d)再对添加的连接基因后续的所有连接基因增加与添加的算子的连接,新增的连接情况是随机的有连接或无连接,其他原有的连接保持不变;e)调用染色体合理性检测函数对添加后的模型染色体进行检测,如果有冲突则重新添加,直到通过检测为止。

8.根据权利要求1所述的基于进化算法的卷积神经网络结构搜索方法,其特征在于,步骤3)中队列管理器Tq开启后针对队列Q中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器Tw对其训练计算适应度时,worker对其训练计算适应度采用的是根据预设的重复次数进行多次重复训练,多次训练的结果得到多个适应度,且最终得到的适应度为所有适应度中指定数量的前几名的平均值。

9.根据权利要求1所述的基于进化算法的卷积神经网络结构搜索方法,其特征在于,步骤3)中队列管理器Tq开启后针对队列Q中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器Tw对其训练计算适应度后,全局开始记录最佳适应度个体的信息表和训练后的模型,在后续的个体训练中,若某个个体的训练结果在解析之后得到的适应度高于当前最佳适应度则将此个体作为新的最佳适应度个体并保存其训练好的模型,同时删除上一个最佳适应度个体的信息和模型;步骤5)输出最佳模型时,则将最终的最佳适应度个体的训练好的模型输出。

10.一种基于进化算法的卷积神经网络结构搜索系统,包括计算机系统,其特征在于:所述计算机系统被编程以执行权利要求1~9中任意一项所述基于进化算法的卷积神经网络结构搜索方法的步骤。

一种基于进化算法的卷积神经网络结构搜索方法及系统

技术领域

[0001] 本发明涉及深度学习的模型设计领域,具体涉及一种基于进化算法的卷积神经网络结构搜索方法及系统,为一种针对给定数据集实现自动建模、调参和训练的解决方案。

背景技术

[0002] 由于计算机软硬件水平的迅速发展,设备的计算能力和存储空间得到很大提升,随之出现了很多的大规模数据集。这些条件使得采用深层卷积神经网络处理大规模图片数据集十分有效,其中出现了诸如AlexNet、VGGNet、ResNet、GoogleNet等优秀的模型,每个模型设定了一种固定的算子组合结构。当前的卷积神经网络主要有卷积、池化、全连接等算子,很多算子还需要再设置一些精确的参数,例如一个卷积算子需要再设定它的卷积核大小、卷积核个数、是否对输入的图片或特征图添加边距等。当一个模型设计完成,意味着其算子组合的结构固定、算子的参数固定。训练时,图片数据传入模型第一个算子,经过计算得到特征图,再将特征图向后层算子传递,再计算得到新的特征图,直到传递到达模型的最后一层,最后的输出结果再传向损失函数,由损失函数计算损失并进行反向传播更新模型各个算子的权值。

[0003] 深层卷积神经网络的这种算子组合模式可以视为一种单向有向无环图,算子为图的节点,特征图在算子之间的单向传递构成一条边。在足够的计算力的支持下,对这样的计算图模型的结构搜索是可行的。

[0004] 在VGGNet、ResNet、GoogleNet等经典网络中,都存在重复子结构这一现象,也就是说一个完整的网络由一个子结构重复堆叠构建而成。这些网络在许多流行的诸如ImageNet、cifar10等大规模图片数据集都表现优秀,这说明了寻找一个优秀的模型可以简化为寻找一个优秀的子结构,然后再对子结构进行组装即可得到一个模型。那么对模型结构进行搜索的搜索空间将大为减少。

[0005] 由于除了多种多样的自然图片外还有与自然图片区别显著的诸多科学图片数据,例如生物医学图片、卫星遥感图片、天文数据等等,即使在许多的数据集上表现优秀的模型也无法保证在任意一个新的数据集上仍然优秀。这意味着,对于一个新的数据集,我们可能需要根据数据集的特点设计合适的模型来训练学习。然而人工设计模型需要极具经验的工程师,设计方法也不能规模化、流程化,且具有不确定性,难免需要费时费力的调参训练。这使得设计一种针对给定数据集实现自动建模、调参和训练的解决方案十分有必要。

[0006] 进化算法是一种启发式的智能搜索算法:设置初始种群后,由适应度函数评估种群中每个个体的适应度,再由选择函数选出具有高适应度的优秀个体,再由进化函数对这些优秀个体进行一系列进化操作得到新的种群,如此不断迭代,种群整体质量不断提升,最终收敛。我们可以发现进化算法与模型搜索的需要十分契合。因此,如何实现基于进化算法的卷积神经网络结构搜索,已经成为一项亟待解决的关键技术问题。

发明内容

[0007] 本发明要解决的技术问题:为了解决人工设计模型不能规模化、流程化、具有不确定性等缺点,提供一种基于进化算法的卷积神经网络结构搜索方法及系统,能针对给定数据集实现自动建模、调参和训练,具有性能高、规模化、流程化、确定性好的优点,尤其适用在高性能计算机集群上部署实施的优点。

[0008] 为了解决上述技术问题,本发明采用的技术方案为:

[0009] 一种基于进化算法的卷积神经网络结构搜索方法,实施步骤包括:

[0010] 1) 输入数据集并设定预设参数;

[0011] 2) 根据预设参数初始化种群,得到初始种群;

[0012] 3) 通过作为主线程的控制器 T_c 将初始种群弹入队列 Q 并开启队列管理器 T_q 和消息管理器 T_m ,队列管理器 T_q 开启后针对队列 Q 中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_w 对其训练计算适应度,且队列管理器 T_q 在完成训练的个体的个数达到预设参数中的总搜索模型个数后终止,且根据worker管理器 T_w 的状态更新字典 W ,字典 W 中记录全部worker的状态且状态为正在训练不可用或空闲可用;消息管理器 T_m 开启后如果收到worker管理器 T_w 发来的训练结果则将其更新字典 M ,且消息管理器 T_m 在收到训练结果达到预设参数中的总搜索模型个数后终止;

[0013] 4) 控制器 T_c 监控字典 M 、队列 Q 和字典 W ;当字典 M 中有足够的可选个体、队列 Q 为空、字典 W 中有指定数量的可用worker,则通过worker进行一次优秀个体选择,选择策略为预设参数中的竞标赛选择策略;然后根据预设参数中的渐进式搜索设置,确定是否要向选择出的优秀个体中添加基因,再根据预设参数中的交配比例、变异概率、染色体分离进化还是组合进化以及基因冲突时的重进化策略,对选择出来的高适应度个体执行进化,进化之后的个体存入队列 Q 中;当控制器 T_c 监控到字典 M 中个体的个数达到预设参数中的总搜索模型个数时,跳转执行步骤5);

[0014] 5) 输出最佳模型,结束。

[0015] 优选地,步骤2)的详细步骤包括:

[0016] 2.1) 检测是否有人工设置的优秀染色体,如果有则读入,再根据预先设置的初始种群大小随机生成额外需要的染色体个数;

[0017] 2.2) 先从可选算子列表 O 中随机选择出设定数量的算子组成算子染色体,再根据预设的策略生成连接基因,再组成染色体,连接染色体生成后与算子染色体通过一个染色体融合函数即可融合为完整的模型染色体,得到初始种群;

[0018] 2.3) 针对初始种群中每一个染色体进行染色体合理性检测,如果未通过则需要舍弃重新随机生成新的染色体,最终得到通过染色体合理性检测的初始种群。

[0019] 优选地,步骤3)中队列管理器 T_q 开启后针对队列 Q 中的未训练染色体将其弹出解码的详细步骤包括:首先将染色体中的算子基因解码成预设的可选算子列表 O 中的算子名称,再根据连接基因将这些算子名称存入一个多层字典,最后再实现重复堆叠的组装过程,得到完整模型对应的一个字典;所述可选算子列表 O 存储已经被实现了的可以直接调用的每一个算子的算子名称,每个算子的类型和参数已经限定,且同一类型的两个算子如果有任何的参数不同,那么在可选算子列表 O 中两者会作为两个不同的算子存放;可选算子列表 O 确定且其中的算子均实现可解码后,每次搜索时每个个体中的全部算子均为可选算子列

表0中的算子,同时每个个体中的算子的总个数也被限定;然后,根据预先设定的第三方模型训练工具遍历存储完整模型的字典,依次调用其中算子针对设定的训练工具的解码实现,得到在训练工具中此算子的字符串描述,根据模型字典将所有算子的字符串描述进行拼接,得到这个工具中模型的描述代码或者描述配置,再将其存储为训练工具可以读入的文件作为worker管理器 T_w 的输入。

[0020] 优选地,步骤3)中worker管理器 T_w 对其训练计算适应度的详细步骤包括:

[0021] 3.1) 检查字典W中是否有空闲可用状态的worker,当且仅当字典W中有空闲可用状态的worker时跳转执行下一步;否则,继续检查字典W;

[0022] 3.2) 选择worker,更新字典W将选中worker的状态设置为训练不可用;

[0023] 3.3) 将解码后得到的模型提交到选中worker进行训练,并检测是否训练结束,当且仅当选中worker训练结束时跳转执行下一步;

[0024] 3.4) 更新字典W将选中worker的状态设置为空闲可用,释放选中worker;

[0025] 3.5) 解析训练结果,调用如式(1)所示的适应度函数计算得到适应度;

[0026] $f = w_a * a + w_m * m + w_t * t$ (1)

[0027] 式(1)中, f 为适应度, w_a 为训练精度 a 的权重函数, w_m 为模型参数量 m 的权重函数, w_t 为训练用时 t 的权重函数,其中训练精度 a 的权重函数 w_a 的值域为非负,模型参数量 m 的权重函数 w_m 的值域为0或负值,训练用时 t 的权重函数 w_t 的值域为0或负值;

[0028] 3.6) 将适应度发送给消息管理器 T_M 。

[0029] 优选地,步骤4)中通过worker进行一次优秀个体选择时,选择的策略为下述方法之一:①、固定数量策略:每次选择可选群体中固定数目的优秀个体进行进化;②、固定比例策略:每次选择可选群体中固定比例的优秀个体进行进化;③、根据可用worker数确定,有多少worker可用则从可选群体中选择多少个体进化;步骤4)中对选择出来的高适应度个体执行进化时采用的进化操作为双点交叉、多点变异、重进化、分离进化与组合进化中的一种或者多种;其中,双点交叉是指从被选择的个体中随机挑选偶数个个体进行双点交叉操作,挑选比例即为交叉比例,将被挑选出的偶数个个体按每组两个个体分组,对每组两个个体的染色体执行双点交叉操作时,随机选择两个基因位点,两个染色体交换这两个位点间的基因片段;多点变异是指针对每个染色体进行多次变异,且每一次的变异概率相同或者不同,每一次变异时针对每个染色体,随机生成一个概率值作为变异概率,若大于预先设定的变异概率,则再随机选择一个位点对其进行变异,变异指0变为1或1变为0,这称为一次变异;重进化是指未通过染色体合理性检测的染色体可以根据预先设定进行再次进化或者重新进化;分离进化是指根据预先设定的参数选择直接对整条模型染色体进行进化,组合进化是指根据预先设定的参数选择将模型染色体分离成算子染色体和连接染色体分别进化。

[0030] 优选地,步骤3)中队列管理器 T_Q 开启后针对队列Q中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_w 对其训练计算适应度后还包括进行早停策略控制步骤,所述早停策略控制的详细步骤包括:首先判断完成训练的个体的个数是否大于预设的早停迭代次数,如果完成训练的个体的个数大于预设的早停迭代次数、且训练结果中的适应度低于预设的适应度阈值,则队列管理器 T_Q 提前终止、而不需要在完成训练的个体的个数达到预设参数中的总搜索模型个数后终止。

[0031] 优选地,步骤3)中队列管理器 T_Q 开启后针对队列Q中的未训练染色体将其弹出解

码后开启一个作为独立临时线程的worker管理器 T_w 对其训练计算适应度时,worker训练计算适应度采用渐进式搜索策略,所述渐进式搜索策略的详细步骤包括:设定多级搜索数阈值,每个阈值表示已经开启训练的模型的个数,其中较低的阈值对应较小的搜索空间,较大的阈值对应较大的搜索空间,这里搜索空间大小的变化通过染色体中算子基因的个数来控制,使得初始种群在第一级搜索空间中构建并依次训练,之后直到已经开启训练的个体的个数达到了第一级阈值,此时对后续被开启训练的个体的染色体中的算子的个数进行统计,如果不满足第二级搜索空间的数量的要求则对其进行算子添加操作,如此继续下去,直到在最后一级的搜索空间完成搜索;其中,进行算子添加操作的详细步骤包括:a)首先分离出算子染色体和连接染色体;b)然后往算子染色体的一个随机位点随机添加一个算子的基因;c)再在相同位点添加连接基因,其中的连接的有无也是随机的;d)再对添加的连接基因后续的所有连接基因增加与添加的算子的连接,新增的连接情况是随机的有连接或无连接,其他原有的连接保持不变;e)调用染色体合理性检测函数对添加后的模型染色体进行检测,如果有冲突则重新添加,直到通过检测为止。

[0032] 优选地,步骤3)中队列管理器 T_q 开启后针对队列Q中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_w 对其训练计算适应度时,worker对其训练计算适应度采用的是根据预设的重复次数进行多次重复训练,多次训练的结果得到多个适应度,且最终得到的适应度为所有适应度中指定数量的前几名的平均值。

[0033] 优选地,步骤3)中队列管理器 T_q 开启后针对队列Q中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_w 对其训练计算适应度后,全局开始记录最佳适应度个体的信息表和训练后的模型,在后续的个体训练中,若某个个体的训练结果在解析之后得到的适应度高于当前最佳适应度则将此个体作为新的最佳适应度个体并保存其训练好的模型,同时删除上一个最佳适应度个体的信息和模型;步骤5)输出最佳模型时,则将最终的最佳适应度个体的训练好的模型输出。

[0034] 本发明还提供一种基于进化算法的卷积神经网络结构搜索系统,包括计算机系统,所述计算机系统被编程以执行本发明所述基于进化算法的卷积神经网络结构搜索方法的步骤。

[0035] 和现有技术相比,本发明具有下述优点:

[0036] 1、本发明能针对给定数据集实现自动建模、调参和训练,具有规模化、流程化、确定性好的优点。

[0037] 2、本发明通过作为主线程的控制器 T_c 将初始种群弹入队列Q并开启队列管理器 T_q 和消息管理器 T_m ,队列管理器 T_q 开启后针对队列Q中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_w 对其训练计算适应度,通过控制器 T_c 、队列管理器 T_q 、worker管理器 T_w 、消息管理器 T_m 的协同完成基于进化算法的卷积神经网络结构的并行搜索,搜索效率高,且尤其适用在高性能计算机集群上部署实施。

附图说明

[0038] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本发明的实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据提供的附图获得其他的附图。

- [0039] 图1为本发明实施例方法的基本流程示意图。
- [0040] 图2为本发明实施例中控制器T_c的基本流程示意图。
- [0041] 图3为本发明实施例中染色体解码的原理示意图。
- [0042] 图4为本发明实施例中队列管理器T_q的工作流程示意图。
- [0043] 图5为本发明实施例中worker管理器T_w的工作流程示意图。
- [0044] 图6为本发明实施例中消息管理器T_m的工作流程示意图。

具体实施方式

[0045] 如图1所示,本实施例基于进化算法的卷积神经网络结构搜索方法的实施步骤包括:

[0046] 1) 输入数据集并设定预设参数;

[0047] 2) 根据预设参数初始化种群,得到初始种群;

[0048] 3) 如图2所示,通过作为主线程的控制器T_c将初始种群弹入队列Q并开启队列管理器T_q和消息管理器T_m,队列管理器T_q开启后针对队列Q中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器T_w对其训练计算适应度,且队列管理器T_q在完成训练的个体的个数达到预设参数中的总搜索模型个数后终止,且根据worker管理器T_w的状态更新字典W,字典W中记录全部worker(工作者/执行者)的状态且状态为正在训练不可用或空闲可用;消息管理器T_m开启后如果收到worker管理器T_w发来的训练结果则将其更新字典M,且消息管理器T_m在收到训练结果达到预设参数中的总搜索模型个数后终止;

[0049] 4) 如图2所示,控制器T_c监控字典M、队列Q和字典W;当字典M中有足够的可选个体、队列Q为空、字典W中有指定数量的可用worker,则通过worker进行一次优秀个体选择,选择策略为预设参数中的竞标赛选择策略;然后根据预设参数中的渐进式搜索设置,确定是否要向选择出的优秀个体中添加基因,再根据预设参数中的交配比例、变异概率、染色体分离进化还是组合进化以及基因冲突时的重进化策略,对选择出来的高适应度个体执行进化,进化之后的个体存入队列Q中;当控制器T_c监控到字典M中个体的个数达到预设参数中的总搜索模型个数时,跳转执行步骤5);

[0050] 5) 输出最佳模型,结束。

[0051] 本实施例中,步骤1)中设置的预设参数包括数据和环境相关参数、确定搜索空间的相关参数、染色体编码解码相关参数、模型训练相关参数、进化算法相关参数。(1) 数据和环境相关参数包括:a) 数据集的地址;b) 输入的图片数据集的形状;c) 输入的数据集标签的形状;d) 测试集的大小;e) 第三方训练工具的选用参数;f) 异步训练时可用worker数量。(2) 确定搜索空间的相关参数包括:a) 染色体中算子个数;b) 染色体生成时可选算子列表。(3) 染色体编码解码相关参数,包括:a) 初始化种群的策略;b) 人工设计的初始种群的染色体列表;c) 模型中重复染色体描述的子结构的次数;d) 连接各个重复的子结构的连接算子;e) 模型中算子允许输出的特征图最小形状。(4) 模型训练相关参数包括:a) 训练时批处理的数量;b) 模型训练迭代次数;c) 模型早停训练迭代次数;d) 模型重复训练次数;e) 取适应度前多少名的重复训练的结果计算最终适应度;f) 模型早停重复训练次数;g) 取适应度前多少名的早停重复训练的结果计算早停的最终适应度;h) 早停训练的适应度阈值;i) 个体训练任务失败后是否重新训练;j) 最佳模型继续训练迭代次数;k) 其他第三方训练工具训练时

需要的超参数。(5) 进化算法相关参数包括:a) 竞标赛选择策略;b) 竞标赛选择选择开启的阈值;c) 交配比例;d) 变异概率;e) 渐进式搜索设置;f) 适应度时精度、模型参数量、和运行时间的权重函数;g) 初始种群的大小;h) 总搜索模型个数;i) 染色体分离进化还是组合进化;j) 基因冲突时的重进化策略。

[0052] 初始化种群时,读取预设参数中的人工设计的初始种群的染色体列表,根据预设参数中的初始种群的大小确定需要补充生成的染色体的数量,再根据预设参数中的确定搜索空间的相关参数随机生成各个属于搜索空间的染色体,得到初始种群。本实施例中,步骤2) 的详细步骤包括:

[0053] 2.1) 检测是否有人工设置的优秀染色体,如果有则读入,再根据预先设置的初始种群大小随机生成额外需要的染色体个数;

[0054] 2.2) 先从可选算子列表0中随机选择出设定数量的算子组成算子染色体,再根据预设的策略生成连接基因,再组成染色体,连接染色体生成后与算子染色体通过一个染色体融合函数即可融合为完整的模型染色体,得到初始种群;

[0055] 2.3) 针对初始种群中每一个染色体进行染色体合理性检测,如果未通过则需要舍弃重新随机生成新的染色体,最终得到通过染色体合理性检测的初始种群。

[0056] 种群初始时,先检测是否有人工设置的优秀染色体,如果有则读入,再根据预先设置的初始种群大小随机生成额外需要的染色体个数。生成时,先从可选算子列表0中随机选择出设定数量的算子组成算子染色体,再根据以下可选策略生成连接基因,再组成染色体:
(1) 低频策略。即当前连接基因的全部前层算子基因中的某个算子基因与其后层的所有算子基因有连接的频率越低,则它被选为当前连接基因的后继算子基因的一个输入的概率越大;
(2) 高频策略。即当前连接基因的全部前层算子基因中的某个算子基因与其后层的所有算子基因有连接的频率越高,则它被选为当前连接基因的后继算子基因的一个输入的概率越大;
(3) 小尺寸策略。即当前连接基因的全部前层算子基因中的某个算子基因解码为算子后的输出特征图尺寸越小,则它被选为当前连接基因的后继算子基因的一个输入的概率越大;
(4) 公平策略。即当前连接基因的全部前层算子基因中的每个算子被选为当前连接基因的后继算子基因的一个输入的概率均完全随机。连接染色体生成后与算子染色体通过一个染色体融合函数即可融合为完整的模型染色体。这些模型染色体可能存在基因冲突,即解码后不符合算子多特征图输入的的形状一致性要求、最终层输出特征图形状的最低尺寸要求、是否没有输出和是否没有输入。判断是否存在基因冲突的函数称作染色体合理性检测函数。初始种群中每一个染色体均需要通过染色体合理性检测函数的检测,如果未通过则需要舍弃重新随机生成新的染色体。

[0057] 可选算子列表0已经被实现了的可以直接调用的每一个算子的算子名称,每个算子的类型和参数已经限定,意味着同一类型的两个算子如果有任何的参数不同,那么在可选算子列表0中两者会作为两个不同的算子存放。对于卷积和池化算子,过滤器尺寸在可选算子列表0中可以根据需要限定为几个常用的尺寸、步长可以根据需要限定为几个常用的步长,对于卷积算子,过滤器的数量设置也可以根据需要限定为几个常用的设置。可选算子列表0确定且其中的算子均实现可解码后,每次搜索时,每个个体中的全部算子均为0中的算子,同时,每个个体中的算子的总个数也被限定。此外,模型子结构组装为完整模型时,所有的连接算子和输入层算子、分类部分算子和输出层算子均预先设定。由此,整个搜索空间

则被限定。

[0058] 本实施例中,染色体编码方案如下:以可选算子列表0中每个算子的索引作为其在染色体中的基因,然后以列表0的长度转换为二进制所需的位数作为一个固定长度,将这些算子基因按照这个固定长度位数转换为二进制。举例来说,若可选算子列表0中有8个算子,长度转换为二进制为1000,那么所有算子基因均能转换为4位二进制码。算子基因与算子基因之间留出连接基因的空位,空位的长度即为在当前连接基因之前的算子基因的个数。然后空位中每个位用0或1填充,在第几个位置填1则意味着染色体顺序第几个算子基因是当前连接基因的后继算子基因的输入。这样连接基因的后继算子基因通过连接基因与连接基因之前的所有算子基因构成连接关系。一个算子基因的前继连接基因至少有1个1,且有几个1,意味着染色体对应模型中这个算子有几个来自前层算子的输出作为输入。同理,一个算子基因的后继连接基因至少有1个1,且有几个1,意味着染色体对应模型中有几个后层算子将这个算子的输出作为输入。算子基因和连接基因根据上述编码方法构成一条完整的模型染色体,若将一条完整的模型染色体中的算子基因与连接基因通过一个染色体分离函数保持原有顺序分离,则可得到一条算子染色体和一条连接染色体。

[0059] 如图2所示,控制器T_c即主线程,构建初始种群,并将初始种群依次弹入队列Q,Q为存放为未训练的染色体的共享队列;然后,控制器T_c开启队列管理器T_q即一个待训练模型队列的监控线程,和消息管理器T_m即一个训练线程的消息队列的监控线程,共同维护三个共享的字典M、队列Q和字典W,其中M存放已训练的染色体及其适应度,W中为全部worker及其状态(正在训练不可用或空闲可用);T_q开启后监控Q是否为空,只要Q中有一个未训练染色体则弹出该染色体,对其解码后开启一个worker管理器T_w即一个独立的临时线程,负责对其训练计算适应度。T_q记录自己开启训练的个体的个数,达到设定数量则终止线程;任何一个T_w被开启后就会立即检查W中是否有空闲可用状态的worker,有则将线程收到的模型提交到此worker上调用第三方训练工具进行训练并将W中该worker的状态更新为正在训练不可用,没有则继续监控直到有为止。在训练结束后线程将W中该worker的状态更新为空闲可用并将训练结果解析后发送给线程T_m,然后终止线程;T_m接收到任何一个T_w发来的训练结果则根据结果更新M。T_m记录自己接收到的完成训练的个体的个数,达到设定数量则终止线程;T_c监控M、Q和W,一旦M中有足够的可选个体(数量大于设定阈值)、Q为空、W中有一定数量的可用worker(数量大于设定阈值),则进行一次优秀个体选择,再对选择出来的高适应度个体执行进化,进化之后的个体存入Q中。主线程监控M中个体的个数,达到设定数量则终止线程。

[0060] 一条染色体编码的内容实际上是模型的一个子结构,解码时,会根据预先设置的重复次数将这个子结构重复堆叠,组装成一个大的模型,其中子结构与子结构之间的连接算子已预先设定。本实施例中,步骤3)中队列管理器T_q开启后针对队列Q中的未训练染色体将其弹出解码的详细步骤包括:首先将染色体中的算子基因解码成预设的可选算子列表0中的算子名称,再根据连接基因将这些算子名称存入一个多层字典,最后再实现重复堆叠的组装过程,得到完整模型对应的一个字典;然后,根据预先设定的第三方模型训练工具遍历存储完整模型的字典,依次调用其中算子针对设定的训练工具的解码实现,得到在训练工具中此算子的字符串描述,根据模型字典将所有算子的字符串描述进行拼接,得到这个工具中模型的描述代码或者描述配置,再将其存储为训练工具可以读入的文件作为worker

管理器 T_w 的输入。

[0061] 如图3所示,原始的染色体为二进制,二进制的染色体由算子基因和连接基因组成,将二进制的染色体转换为十进制的染色体,将十进制的染色体中的算子基因解码成预设的可选算子列表 O 中的算子名称,再根据连接基因将这些算子名称存入一个多层字典,最后再实现重复堆叠的组装过程,得到完整模型对应的一个字典,得到模型子结构;然后,根据预先设定的第三方模型训练工具遍历存储完整模型的字典,依次调用其中算子针对设定的训练工具的解码实现,得到在训练工具中此算子的字符串描述,根据模型字典将所有算子的字符串描述进行拼接,得到这个工具中模型的描述代码或者描述配置,再将其存储为训练工具可以读入的文件作为worker管理器 T_w 的输入。解码函数首先将染色体中的算子基因解码成可选算子列表 O 中的算子名称,再根据连接基因将这些算子名称存入一个多层字典,最后再实现重复堆叠的组装过程,得到完整模型对应的一个字典。然后,解码函数根据预先设定的第三方模型训练工具,例如caffe、TensorFlow或Pytorch,遍历存储完整模型的字典,依次调用其中算子针对设定的训练工具的解码实现,得到在训练工具中此算子的字符串描述,根据模型字典将所有算子的字符串描述进行拼接,得到这个工具中模型的描述代码或者描述配置,再将其存储为训练工具可以读入的文件,例如caffe模型配置文件与caffe求解器配置文件、TensorFlow代码文件或Pytorch代码文件。

[0062] 如图4所示,队列管理器 T_q 开启后针对队列 Q 中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_w 对其训练计算适应度。初始时开启训练的个体的个数赋值为0,且每弹出一个染色体,则启动一个worker管理器 T_w 对其训练计算适应度,并将开启训练的个体的个数加1,且队列管理器 T_q 在完成训练的个体的个数达到预设参数中的总搜索模型个数后终止。

[0063] 个体的训练通过调用第三方模型训练工具进行训练。以调用caffe为例,解码函数根据模型染色体的编码和算子的caffe解码实现,解码并组装出caffe的模型配置文件,根据预设的其他参数生成caffe的求解器配置文件,再调用caffe对个体进行训练,每个个体训练的迭代次数、学习率等超参数为相同的设定值。训练结束对输出文件进行解析得到训练的精度、用时。模型参数量通过模型各个算子的参数设置可以直接计算。一个个体的适应度为其训练精度、模型参数量和训练用时的加权和。

[0064] 如图5所示,步骤3)中worker管理器 T_w 对其训练计算适应度的详细步骤包括:

[0065] 3.1) 检查字典 W 中是否有空闲可用状态的worker,当且仅当字典 W 中有空闲可用状态的worker时跳转执行下一步;否则,继续检查字典 W ;

[0066] 3.2) 选择worker,更新字典 W 将选中worker的状态设置为训练不可用;

[0067] 3.3) 将解码后得到的模型提交到选中worker进行训练,并检测是否训练结束,当且仅当选中worker训练结束时跳转执行下一步;

[0068] 3.4) 更新字典 W 将选中worker的状态设置为空闲可用,释放选中worker;

[0069] 3.5) 解析训练结果,调用如式(1)所示的适应度函数计算得到适应度;

[0070]
$$f = w_a * a + w_m * m + w_t * t \quad (1)$$

[0071] 式(1)中, f 为适应度, w_a 为训练精度 a 的权重函数, w_m 为模型参数量 m 的权重函数, w_t 为训练用时 t 的权重函数,其中训练精度 a 的权重函数 w_a 的值域为非负,模型参数量 m 的权重函数 w_m 的值域为0或负值,训练用时 t 的权重函数 w_t 的值域为0或负值;即:

[0072] $w_a > 0, w_m \leq 0, w_t \leq 0$;

[0073] 3.6) 将适应度发送给消息管理器 T_M 。

[0074] 如图6所示,消息管理器 T_M 开启后如果收到worker管理器 T_W 发来的训练结果则将其更新字典 M ,且消息管理器 T_M 在收到训练结果达到预设参数中的总搜索模型个数后终止。

[0075] 本实施例中,步骤4)中通过worker进行一次优秀个体选择时,选择的策略为下述方法之一:①、固定数量策略:每次选择可选群体中固定数目的优秀个体进行进化;②、固定比例策略:每次选择可选群体中固定比例的优秀个体进行进化;③、根据可用worker数确定,有多少worker可用则从可选群体中选择多少个体进化。

[0076] 本实施例中,步骤4)中对选择出来的高适应度个体执行进化时采用的进化操作作为双点交叉、多点变异、重进化、分离进化与组合进化中的一种或者多种;其中,双点交叉是指从被选择的个体中随机挑选偶数个个体进行双点交叉操作,挑选比例即为交叉比例,将被挑选出的偶数个个体按每组两个个体分组,对每组两个个体的染色体执行双点交叉操作时,随机选择两个基因位点,两个染色体交换这两个位点间的基因片段;多点变异是指针对每个染色体进行多次变异,且每一次的变异概率相同或者不同,每一次变异时针对每个染色体,随机生成一个概率值作为变异概率,若大于预先设定的变异概率,则再随机选择一个位点对其进行变异,变异指0变为1或1变为0,这称为一次变异;重进化是指未通过染色体合理性检测的染色体可以根据预先设定进行再次进化或者重新进化;分离进化是指根据预先设定的参数选择直接对整条模型染色体进行进化,组合进化是指根据预先设定的参数选择将模型染色体分离成算子染色体和连接染色体分别进化。

[0077] 本实施例包含早停策略,具体策略如下:步骤3)中队列管理器 T_Q 开启后针对队列 Q 中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_W 对其训练计算适应度后还包括进行早停策略控制步骤,所述早停策略控制的详细步骤包括:首先判断完成训练的个体的个数是否大于预设的早停迭代次数,如果完成训练的个体的个数大于预设的早停迭代次数、且训练结果中的适应度低于预设的适应度阈值,则队列管理器 T_Q 提前终止、而不需要在完成训练的个体的个数达到预设参数中的总搜索模型个数后终止。通过设定适应度阈值和早停迭代次数,worker管理器 T_W 中先进行一个迭代次数为早停迭代次数的预训练,预训练的适应度结果如果低于适应度阈值,则直接以预训练结果为最终结果,不再进行完整训练,如果预训练结果高于适应度阈值,则再继续进行完整训练,以完整训练结果作为最终结果,该策略的使用与否是可选的。通过该策略,能够提升训练效率、减少时间花费,能够实现加速搜索。

[0078] 本实施例包含渐进式搜索策略,具体策略如下:步骤3)中队列管理器 T_Q 开启后针对队列 Q 中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_W 对其训练计算适应度时,worker训练计算适应度采用渐进式搜索策略,所述渐进式搜索策略的详细步骤包括:设定多级搜索数阈值,每个阈值表示已经开启训练的模型的个数,其中较低的阈值对应较小的搜索空间,较大的阈值对应较大的搜索空间,这里搜索空间大小的变化通过染色体中算子基因的个数来控制,使得初始种群在第一级搜索空间中构建并依次训练,之后直到已经开启训练的个体的个数达到了第一级阈值,此时对后续被开启训练的个体的染色体中的算子的个数进行统计,如果不满足第二级搜索空间的数量要求则对其进行算子添加操作,如此继续下去,直到在最后一级的搜索空间完成搜索;其中,进行算子添

加操作的详细步骤包括:a)首先分离出算子染色体和连接染色体;b)然后往算子染色体的一个随机位点随机添加一个算子的基因;c)再在相同位点添加连接基因,其中的连接的有无也是随机的;d)再对添加的连接基因后续的所有连接基因增加与添加的算子的连接,新增的连接情况是随机的有连接或无连接,其他原有的连接保持不变;e)调用染色体合理性检测函数对添加后的模型染色体进行检测,如果有冲突则重新添加,直到通过检测为止。通过该策略,能够提升训练效率、减少时间花费,能够实现加速搜索。

[0079] 本实施例包含多次平均策略,具体策略如下:步骤3)中队列管理器 T_Q 开启后针对队列 Q 中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_W 对其训练计算适应度时,worker对其训练计算适应度采用的是根据预设的重复次数进行多次重复训练,多次训练的结果得到多个适应度,且最终得到的适应度为所有适应度中指定数量的前几名的平均值,具体前几名为预先设定,该策略的使用与否是可选的,通过该策略即在上述的早停训练和完整训练中,将单次训练改为多次重复训练,早停训练的重复次数和完整训练的重复次数预先设定,从而能够提高适应度的计算准确度。

[0080] 本实施例包含最佳模型输出策略,具体策略如下:步骤3)中队列管理器 T_Q 开启后针对队列 Q 中的未训练染色体将其弹出解码后开启一个作为独立临时线程的worker管理器 T_W 对其训练计算适应度后,全局开始记录最佳适应度个体的信息表和训练后的模型,在后续的个体训练中,若某个个体的训练结果在解析之后得到的适应度高于当前最佳适应度则将此个体作为新的最佳适应度个体并保存其训练好的模型,同时删除上一个最佳适应度个体的信息和模型;这样,全局只用保存一个训练好的模型,其他模型训练后不必保存,此最佳模型不断更新。在完成搜索后需要输出全部训练个体的最佳模型,此时,根据预先设定,如果还需要继续补充训练,则将重新读入该最佳模型进行预定迭代次数的继续训练,在完成训练后保存模型并输出,此为本次搜索方案最终的输出;如果不需要补充训练则当前最佳模型即为本次搜索方案最终的输出。因此,步骤5)输出最佳模型时,则将最终的最佳适应度个体的训练好的模型输出:根据预设参数中的最佳模型继续训练迭代次数,确定是否对搜索中保存的最佳模型继续训练以及继续训练的迭代次数,然后得到最终训练好的最佳模型。

[0081] 本实施例还提供一种基于进化算法的卷积神经网络结构搜索系统,包括计算机系统,计算机系统被编程以执行本实施例前述基于进化算法的卷积神经网络结构搜索方法的步骤,该计算机系统优选为高性能计算机集群系统。

[0082] 以上所述仅是本发明的优选实施方式,本发明的保护范围并不仅限于上述实施例,凡属于本发明思路下的技术方案均属于本发明的保护范围。应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明原理前提下的若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

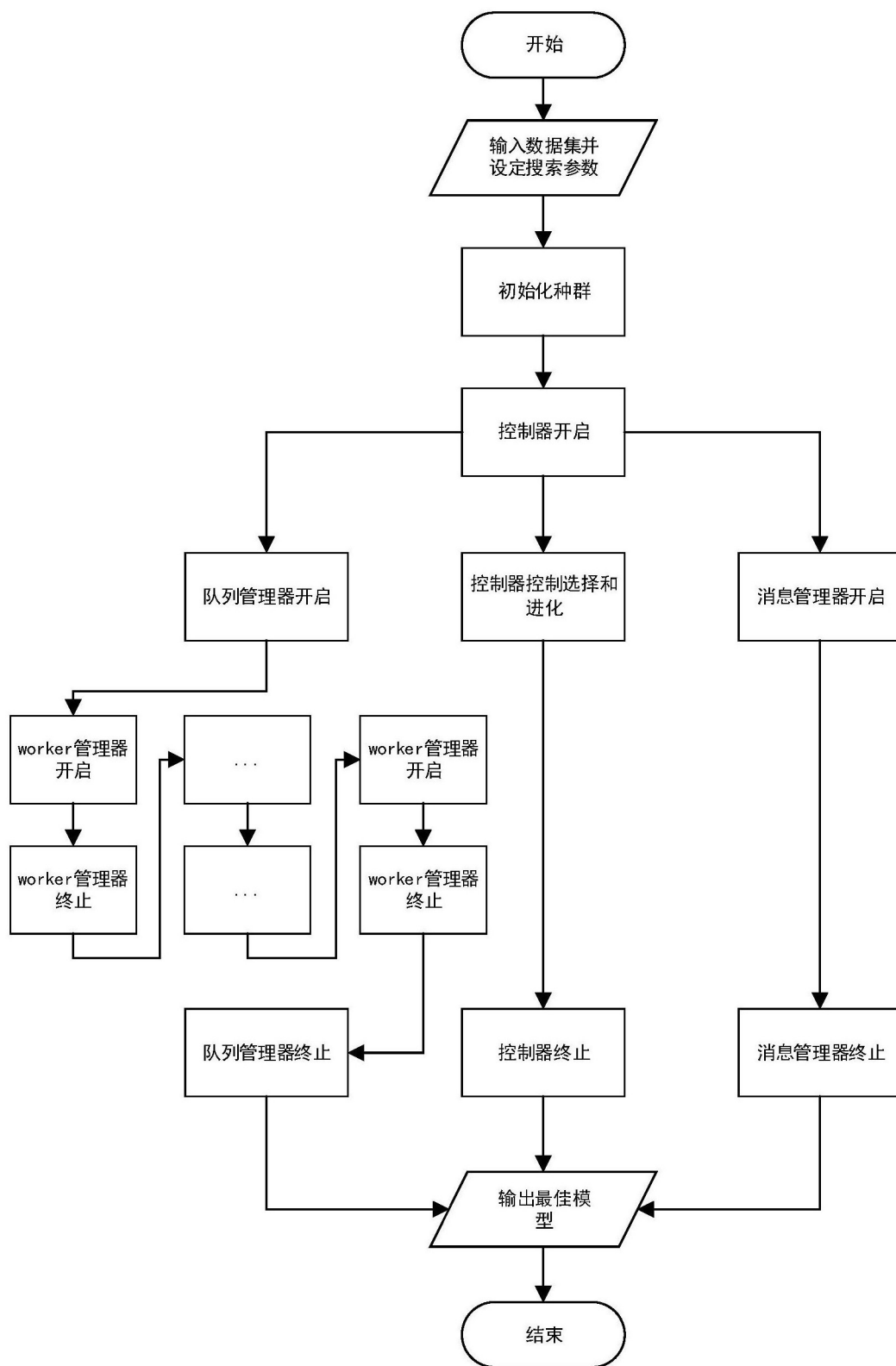


图1

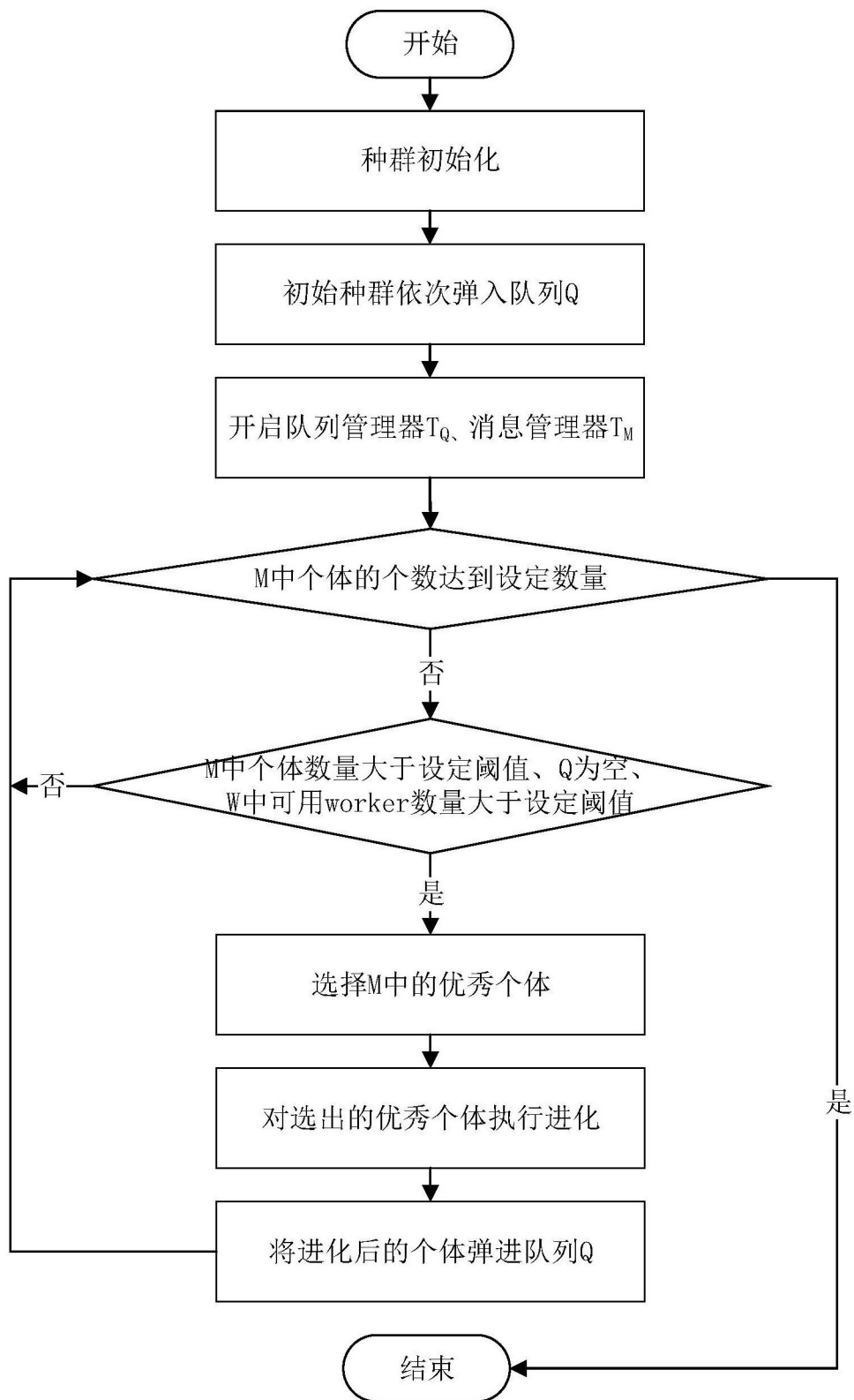


图2

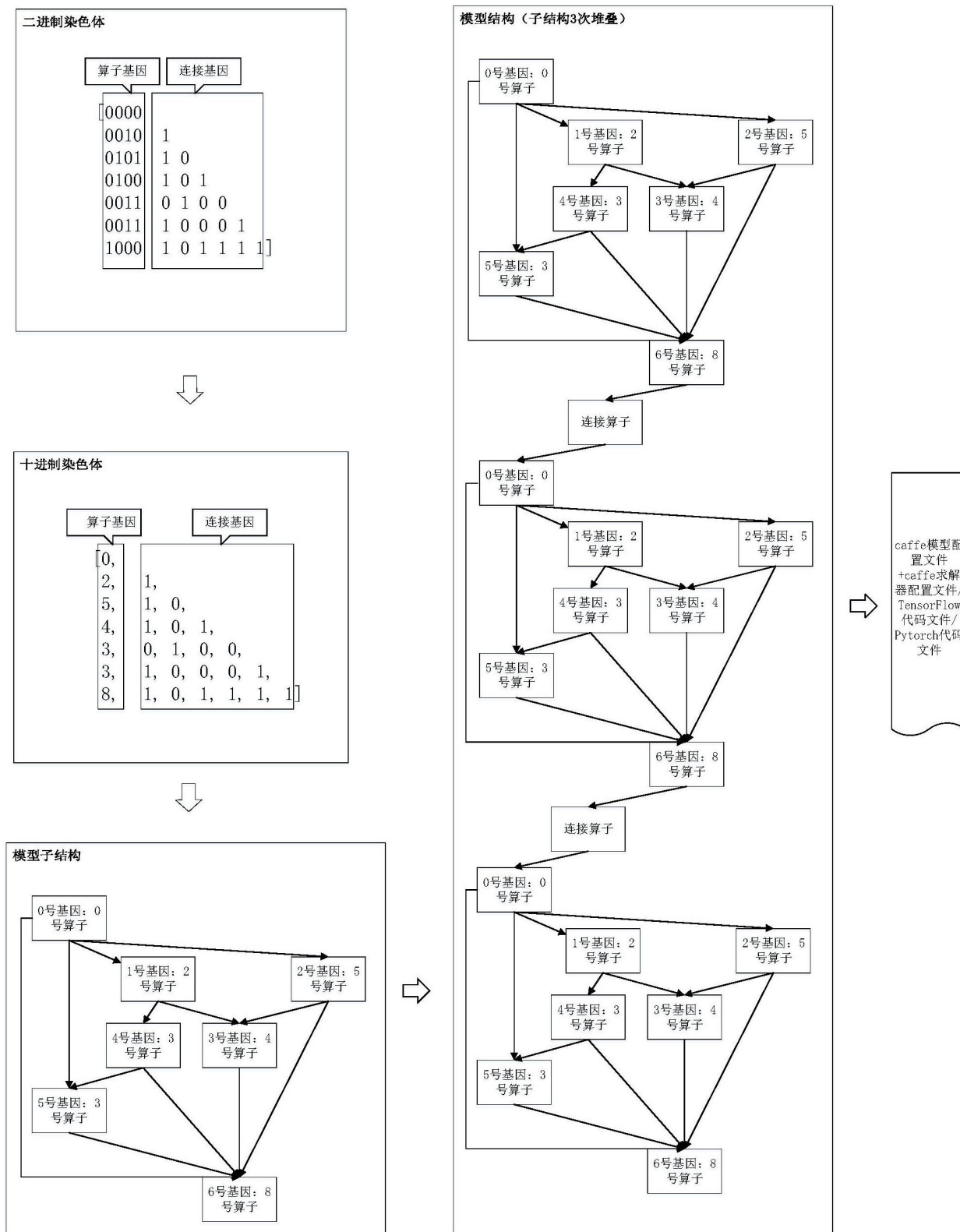


图3

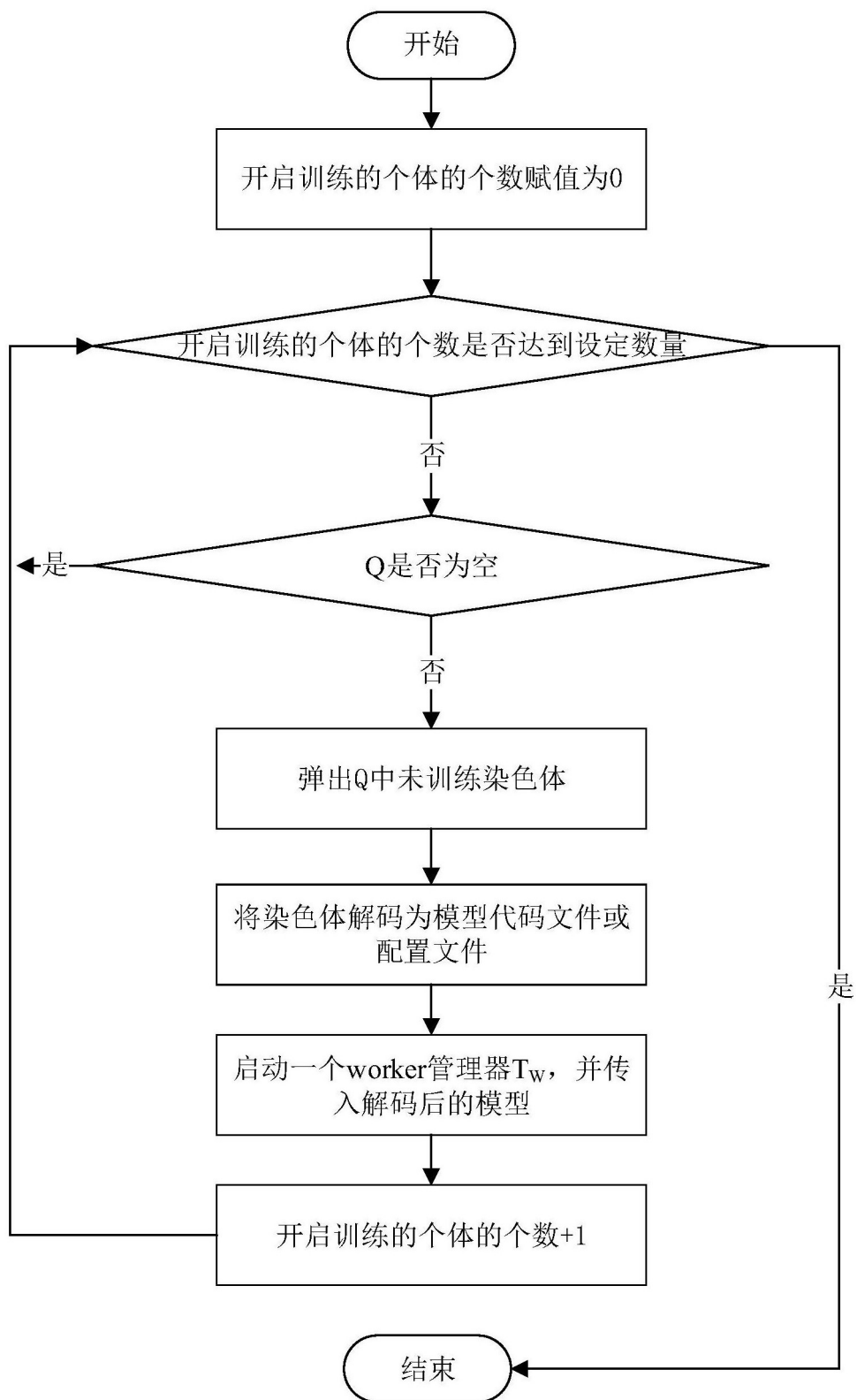


图4

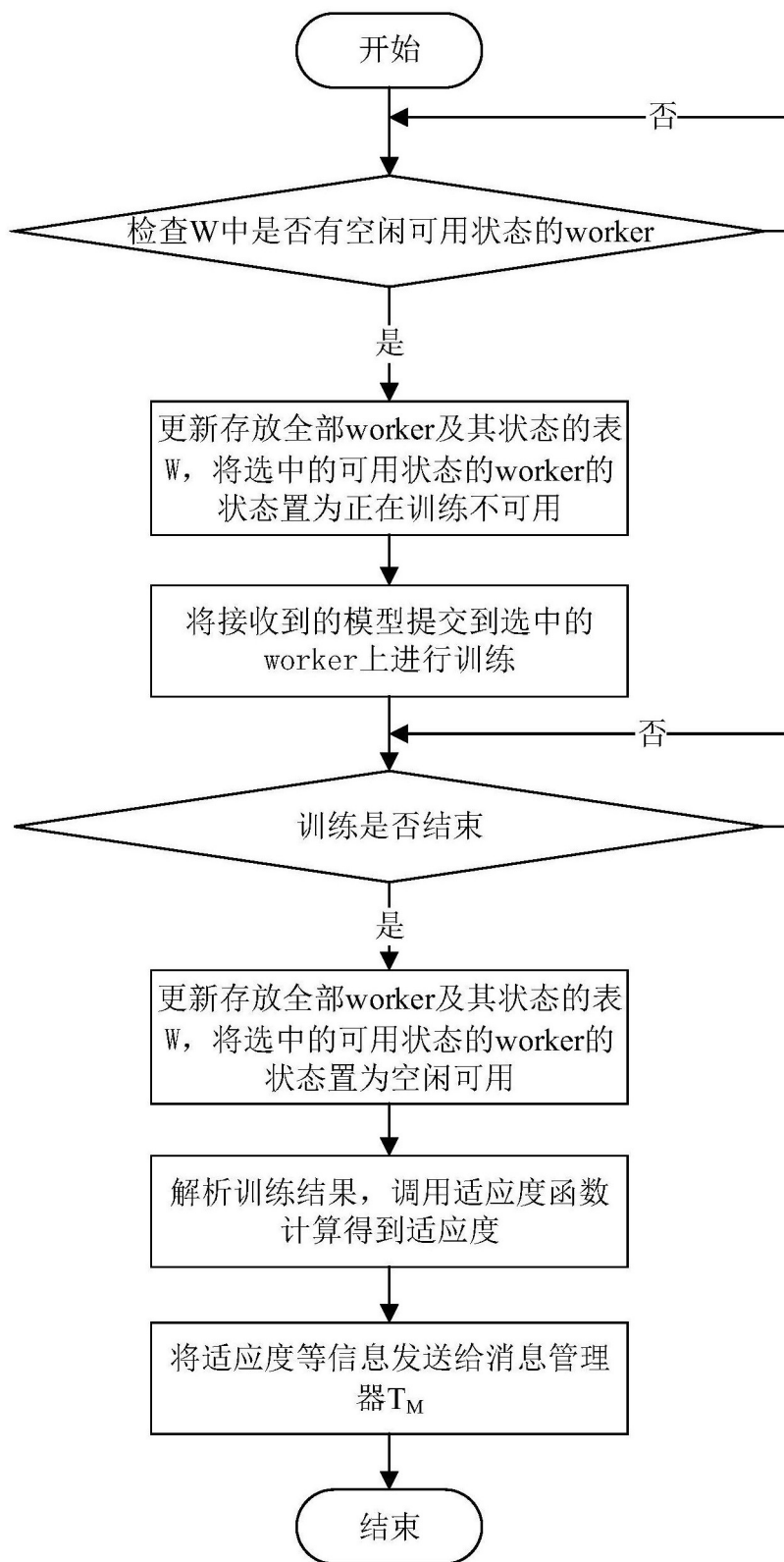


图5

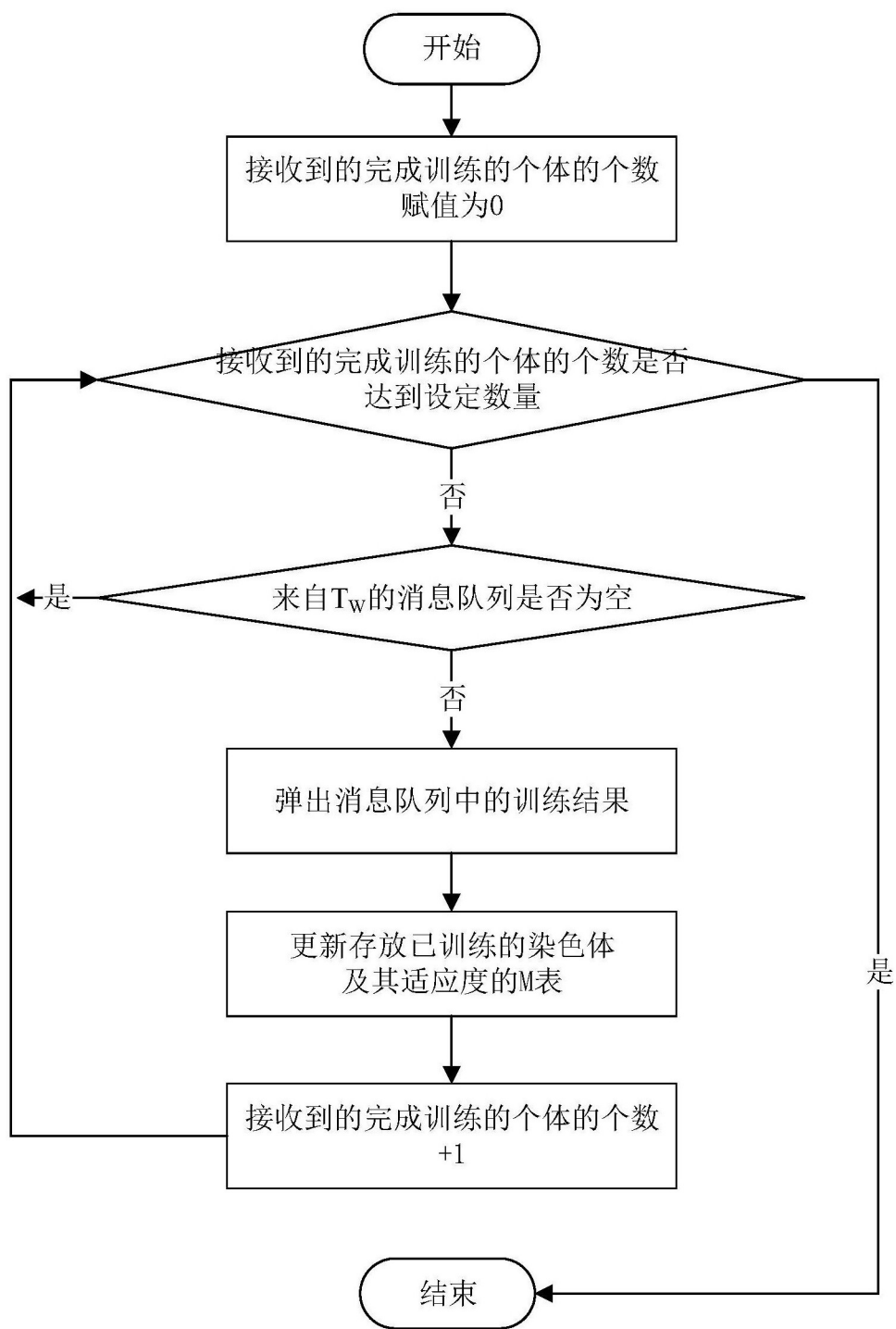


图6