(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
2023 年 6 月 22 日 (22.06.2023)

WIPO | PCT

(10) 国际公布号
WO 2023/108317 A1

(51) 国际专利分类号:
G06K 9/62 (2022.01)

(21) 国际申请号: PCT/CN2021/137348

(22) 国际申请日: 2021 年 12 月 13 日 (13.12.2021)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 中国科学院深圳先进技术研究院 (**SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY CHINESE ACADEMY OF SCIENCES**) [CN/CN]; 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。

(72) 发明人: 周冬豪 (**ZHOU, Donghao**); 中国广东省深圳市南山区深圳大学城学苑大道 1068 号, Guangdong 518055 (CN)。 陈广勇 (**CHEN, Guangyong**); 中国广东省深圳市南山区深圳大学城学苑大道1068号, Guangdong 518055 (CN)。

(74) 代理人: 北京市诚辉律师事务所 (**BEIJING CHENGHUI LAW FIRM**); 中国北京市朝阳区朝阳北路99号楼2单元905, Beijing 100123 (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, 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))。

(54) **Title:** DEEP MULTI-LABEL CLASSIFICATION NETWORK ROBUST TRAINING METHOD BASED ON INFORMATION ENTROPY MAXIMIZATION REGULARIZATION MECHANISM

(54) 发明名称: 基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法

(57) **Abstract:** Disclosed in the present invention is a deep multi-label classification network robust training method based on an information entropy maximization regularization mechanism, which method aims to solve the problem of robust learning of a deep multi-label classification network when each sample only has a single positive label and the remaining labels are missing. An SPML problem is solved by using a binary cross-entropy loss function regularization mechanism based on information entropy maximization, and the aim is to maximize an information entropy of a prediction probability of the deep multi-label classification network for an unknown label, such that a model can avoid the influence of false negative label noise; and learning is performed by using supervision information provided by a true positive label, and a prediction having relatively strong distinguishability is performed on the unknown label. The method is easily implemented, does not introduce any additional learnable parameter, and can be combined with any deep multi-label classification network, thus being suitable for processing a super-large-scale multi-label image data set.

(57) 摘要: 本发明公开了基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，旨在解决每个样本仅有单个正标签且其余标签均缺失的情况下，深度多标签分类网络的鲁棒学习问题，利用基于信息熵最大化的二元交叉熵损失函数正则化机制解决SPML问题，目的是最大化深度多标签分类网络对于未知标签预测概率的信息熵，使得模型可以免受假阴性标签噪声的影响，利用真实正标签提供的监督信息进行学习，并对未知标签做出可区分性较强的预测，方法易于实现，不会引入额外的可学习参数，同时能够与任意的深度多标签分类网络相结合，适合处理超大规模多标签图像数据集。

WO 2023/108317 A1

基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法

技术领域

本发明涉及机器学习技术领域，具体涉及基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法。

背景技术

不同于单标签图像分类，多标签图像分类的目的在于为一个图像预测所有对应的标签。因为一个自然场景中天然地包含多个物体和概念，所以多标签图像分类更加符合真实世界中的设定。然而，对大规模多标签图像数据集进行详尽的标签标注十分费事费力，标注所消耗的人力物力成本也远远大于对单标签图像数据集进行标注的情况，因为一个图像对应的潜在的标签数目可能很多，而且较小物体和较稀有物体往往为标注人员所忽略。因此，现存的开源的大规模多标签图像数据集被广泛认为是缺失部分标签的，这促进了学界对于缺失标签情况下的多标签学习问题的研究。考虑到一种极端缺失标签的情况，即每个样本仅有单个正标签被标注，且其余标签均未进行标注，这被称为单正标签情况下的多标签学习（Single Positive Multi-label Learning，SPML）问题。每个样本仅有单个正标签的设定符合多数实际应用场景，例如从互联网通过查询的方式进行图像数据集的收集等，这表明 SPML 问题具有广泛的实际应用价值。此外，深入研究 SPML 问题并提出妥善的解决方法，可以放宽大规模多标签图像数据集的标注要求，显著降低其标注成本，具有重大的社会经济效益。因此，SPML 是一种值得进一步探索的多标签学习变种，并成为了近年来深度学习领域颇具挑战性的新兴研究方向之一。

相关技术广泛采用的深度多标签分类网络的学习范式是将多个标签的预测问题转化为多个单独的二分类问题。由于多标签设定的特殊性，在大规模多标签图像数据集中，负标签的数量远远大于正标签的数量。因此，一种直观的方法是将所有未知的标签均视作负标签，再利用标准的二元交叉熵损失函数进行训练，这被称为 Assuming Negatives (AN)，它是解决 SPML 问题的基线方法。假设给定一个仅有单个正标签的样本对$(\mathbf{x}^{(n)}, \mathbf{z}^{(n)})$，其中$\mathbf{x}^{(n)}$为第$n$个输入图像，$\mathbf{z}^{(n)}$为第$n$个图像对应的标签向量。则 AN 方法的损失函数为：

$$L_{\mathrm{AN}}(\mathbf{f}^{(n)}, \mathbf{z}^{(n)}) = -\frac{1}{C}\sum_{c=1}^{C}[1_{[z_c^{(n)}=1]}\log(f_c^{(n)}) + 1_{[z_c^{(n)}=0]}\log(1-f_c^{(n)})],$$

其中，$\mathbf{f}^{(n)}$是模型关于$\mathbf{x}^{(n)}$的输出，$f_c^{(n)}$是模型在第c个类上的标签预测概率。C是类别数目，$z_c^{(n)}=1$和$z_c^{(n)}=0$分别表示第c个类存在正标签和负标签。AN 方法的损失函数沿袭了标

准的二元交叉熵损失函数，基于“将所有未知的标签均视作负标签”的标签假设条件进行深度多标签分类网络的训练。

为了解决深度多标签分类网络的训练过程中缺少负标签会陷入退化解的问题，正则化在线标签估计（Regularized Online Label Estimation，ROLE）方法利用数据集中每个样本的平均真实正标签数量来对深度多标签分类网络的输出进行约束，使其模型关于 $\mathbf{x}^{(n)}$ 的输出 $\mathbf{f}^{(n)}$ 的总和不至于过大，这相当于给损失函数引入了一个正则化项，称为期望正标签正则化（Expected Positive Regularization，EPR）。为了对未知标签做出预测以提供更多的监督信息，ROLE 方法在原始的深度多标签分类网络上额外添加了一个标签估计模块，采用互相监督的方式将两者进行联合训练，利用标签估计模块估计出来的标签作为深度多标签分类网络计算损失时的标签，反之，利用深度多标签分类网络的预测作为标签估计模块估计标签时的监督信息。同样地，ROLE 方法也给标签估计模块引入了 EPR 用于约束其估计出来的标签。在进行预测时，ROLE 方法可以丢弃之前训练的标签估计模块，仅用训练完毕的深度多标签分类网络进行标签预测。

然而相关技术具有以下缺点：1）虽然负标签在大规模多标签图像数据集的标签注释中占绝大多数，但是 AN 方法做出的假设（即把所有未知的标签均视作负标签）会引入大量的假阴性标签噪声，这给深度多标签分类网络的训练提供了大量的错误的监督信息，使其难以做出较为正确的正负标签预测。此外，该假设还会加强多标签学习中的正负标签不均衡的现象，使得网络更加倾向于预测出负标签，这进一步削弱了深度多标签分类网络的分类性能。2）ROLE 方法引入的 EPR 正则化方法需要提前统计数据集中每个样本的平均真实正标签数量，这在实际应用场景中是不现实的，因为数据集中真实正标签的数目本来就是未知的。此外，ROLE 方法采用的标签估计模块本质上是一个巨大的估计标签矩阵，这会给原本的网络模型引入额外的大量的可学习参数，大大增加网络训练时的时间和空间需求，且引入的可学习参数数目与数据集中的类别数目和样本数目正相关，这表明 ROLE 方法难以处理超大规模多标签图像数据集。

发明内容

为了解决现有技术中的问题，本发明提出基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，利用基于信息熵最大化的二元交叉熵损失函数正则化机制解决SPML问题，目的是最大化深度多标签分类网络对于未知标签预测概率的信息熵，使得模型可以免受假阴性标签噪声的影响，利用真实正标签提供的监督信息进行学习，并对未知标签做出可区分性较强的预测，方法易于实现，不会引入额外的可学习参数，同时能够与任意的深度多标

签分类网络相结合，适宜处理超大规模多标签图像数据集。

为了实现以上目的，本发明提供了基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，包括：深度多标签分类网络对于每一个样本的每一个未知标签的预测均视为一个离散随机变量，包括预测为正标签和预测为负标签两种，预测为正标签的概率即为深度多标签分类网络输出的未知标签预测概率，控制深度多标签分类网络输出的未知标签预测概率的信息熵最大化。

进一步地，所述方法中假设给定一个仅有单个正标签的样本对$(\mathbf{x}^{(n)}, \mathbf{z}^{(n)})$，其中$\mathbf{x}^{(n)}$为第$n$个输入图像，$\mathbf{z}^{(n)}$为第$n$个图像对应的标签向量，则损失函数为：

$$L_{\mathrm{EMR}}(\mathbf{f}^{(n)}, \mathbf{z}^{(n)}) = -\frac{1}{C}\sum_{c=1}^{C}[1_{[z_c^{(n)}=1]}\log(f_c^{(n)}) + 1_{[z_c^{(n)}=0]}R(f_c^{(n)})]$$

其中，$\mathbf{f}^{(n)}$是模型关于$\mathbf{x}^{(n)}$的输出，$f_c^{(n)}$是模型在第c个类上的预测概率，C是类别数目，$z_c^{(n)}=1$表示第c个类存在正标签，$z_c^{(n)}=0$表示第c个类存在负标签，$R(f_c^{(n)})$为正则化项。

进一步地，所述正则化项$R(f_c^{(n)})$定义为：

$$R(f_c^{(n)}) = -\alpha[f_c^{(n)}\log(f_c^{(n)}) + (1-f_c^{(n)})\log(1-f_c^{(n)})]$$

其中，α为控制正则化强度大小的超参数。

进一步地，所述超参数α的值在 0 到 1 之间进行调节。

进一步地，所述方法中预测为正标签的概率即为网络输出的未知标签预测概率$f_c^{(n)}$，预测为负标签的概率即为$1-f_c^{(n)}$。

进一步地，所述方法会控制未知标签预测概率趋于 0.5。

进一步地，所述方法中假设有一个离散随机变量$X \in \{x_1, x_2, ..., x_n\}$，则其信息熵的定义为：

$$H(X) = \mathrm{E}[I(X)] = \mathrm{E}[-\ln(P(X))]$$

其中，$P(X)$为X的概率质量函数，$E[\cdot]$为期望函数，$I(X)$即为X的信息量。

进一步地，所述方法中若已知或已对$P(X)$做出估计后，则信息熵的公式表示为：

$$H(X) = \sum_i P(x_i)I(x_i) = -\sum_i P(x_i)\log_b P(x_i)$$

其中，b是对数所使用的底。

进一步地，所述方法中当$b=2$，信息熵的单位是 bit；当$b=e$，信息熵的单位是 nat；当$b=10$，信息熵的单位是 Hart。

进一步地，所述方法中深度多标签分类网络对于每一个样本的每一个未知标签的预测均视为一个离散随机变量$X=\{x_1, x_2\}$，事件x_1即为预测为正标签，事件x_2即为预测为负标签，

则 $P(x_1)$ 即为网络输出的未知标签预测概率， $P(x_2)=1-P(x_1)$ 即为预测为负标签的概率。

与现有技术相比，本发明旨在解决每个样本仅有单个正标签且其余标签均缺失的情况下，深度多标签分类网络的鲁棒学习问题，基于信息熵最大化的二元交叉熵损失函数正则化机制，即 EMR 方法，从一个独特的角度来看待数据集中的未知标签，即不对这些未知标签做出任何假设并且承认它们是未知的，在模型训练时的损失函数中最大化深度多标签分类网络对于未知标签预测概率的信息熵，用于解决 SPML 问题，使得模型可以免受假阴性标签噪声的影响，利用真实正标签提供的监督信息进行学习，并对未知标签做出可区分性较强的预测。EMR 方法易于实现，不会引入额外的可学习参数，同时能够与任意的深度多标签分类网络相结合，性能优异，适宜处理超大规模多标签图像数据集。利用本发明所提出的方法进行训练，深度多标签分类网络的泛化性能可以大幅提高，在极度标签缺失的情况下表现良好。

具体实施方式

为使本发明实施例的目的、技术方案和优点更加清楚，下面将对本发明实施例中的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。因此，本发明的实施例的详细描述并非旨在限制要求保护的本发明的范围，而是仅仅表示本发明的选定实施例。基于本发明中的实施例，本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

本发明提供了基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，包括：深度多标签分类网络对于每一个样本的每一个未知标签的预测均视为一个离散随机变量，包括预测为正标签和预测为负标签，预测为正标签的概率即为深度多标签分类网络输出的未知标签预测概率，控制深度多标签分类网络输出的未知标签预测概率的信息熵最大化。

在信息论中，信息熵指的是接收的每条消息中包含的信息的平均量，又被称为信源熵、平均自信息量等，它可以用于度量消息的不确定性。这里的“消息”代表来自分布或数据流中的事件、样本或特征，在本发明所提出的方法中，其代指深度多标签分类网络输出的未知标签预测概率。熵的单位通常为bit，但也用Sh、nat、Hart等单位进行计量，这取决于熵的定义用到对数的底数。假设有一离散随机变量 $X\in\{x_1,x_2,...,x_n\}$ ，则其信息熵的定义为：

$$H(X)=\mathrm{E}[I(X)]=\mathrm{E}[-\ln(P(X))]$$

其中， $P(X)$ 为 X 的概率质量函数， $E[\cdot]$ 为期望函数， $I(X)$ 即为 X 的信息量。

若已知或已对 $P(X)$ 做出估计后，则信息熵的公式可以表示为：

$$H(X)=\sum_i P(x_i)I(x_i)=-\sum_i P(x_i)\log_b P(x_i)$$

其中，b是对数所使用的底。当$b=2$，熵的单位是bit；当$b=e$，熵的单位是nat；而当$b=10$，熵的单位是 Hart。在本发明的方法中，深度多标签分类网络对于每一个样本的每一个未知标签的预测均可以视为一个离散随机变量$X=\{x_1,x_2\}$，事件x_1即为将其预测为正标签，事件x_2即为将其预测为负标签。则$P(x_1)$即为网络输出的未知标签预测概率，相对应地，预测为负标签的概率即为$P(x_2)=1-P(x_1)$。基于此设定，本发明可以进一步完成基于信息熵最大化的二元交叉熵损失函数正则化机制设计。

为了避免错误的假设带来的假阴性标签噪声的影响，同时为了防止深度多标签分类网络过拟合到大量的负标签，本发明采用了一种朴素但有效的思想来看待数据集中的未知标签，即不对这些未知标签做出任何假设并且承认它们是未知的。基于这种思想，深度多标签分类网络对于未知标签的预测概率应该尽量靠近0.5，即趋于0.5，即预测概率的信息熵应该尽量大。因此，本发明基于信息熵最大化的二元交叉熵损失函数正则化机制用于解决SPML问题，目的是最大化深度多标签分类网络对于未知标签预测概率的信息熵，称为（Entropy Maximization Regularizer，EMR）方法。假设给定一个仅有单个正标签的样本对$(\mathbf{x}^{(n)},\mathbf{z}^{(n)})$，其中$\mathbf{x}^{(n)}$为第$n$个输入图像，$\mathbf{z}^{(n)}$为第$n$个图像对应的标签向量。则EMR方法的损失函数为：

$$L_{\mathrm{EMR}}(\mathbf{f}^{(n)},\mathbf{z}^{(n)})=-\frac{1}{C}\sum_{c=1}^{C}[1_{[z_c^{(n)}=1]}\log(f_c^{(n)})+1_{[z_c^{(n)}=0]}R(f_c^{(n)})]$$

其中，$\mathbf{f}^{(n)}$是模型关于$\mathbf{x}^{(n)}$的输出，$f_c^{(n)}$是模型在第c个类上的预测概率。C是类别数目，$z_c^{(n)}=1$表示第c个类存在正标签，$z_c^{(n)}=0$表示第c个类存在负标签，$R(f_c^{(n)})$为EMR方法引入的正则化项，其定义如下所示：

$$R(f_c^{(n)})=-\alpha[f_c^{(n)}\log(f_c^{(n)})+(1-f_c^{(n)})\log(1-f_c^{(n)})]$$

其中，α为控制正则化强度大小的超参数，其值一般在 0 到 1 之间进行调节。在 EMR 方法中，深度多标签分类网络对于每一个样本的每一个未知标签的预测均可以视为一个离散随机变量，分别有预测为正标签和预测为负标签两种可能性。预测为正标签的概率即为网络输出的未知标签预测概率$f_c^{(n)}$，相对应的，预测为负标签的概率即为$1-f_c^{(n)}$。这样一来，EMR方法可以给未知标签施予最大信息熵的监督，而不是把它们全部当作负标签来处理，这契合了本发明采用的思想。EMR 方法使得深度多标签分类网络可以利用真实正标签提供的监督信息进行学习，并对未知标签做出可区分性较强的预测，而不是对它们做出置信度过高甚至是错误的判断，这大幅提升了网络的泛化能力。

本发明所提出的基于信息熵最大化的二元交叉熵损失函数正则化机制，即 EMR 方法，

从一个独特的角度来看待数据集中的未知标签，即不对这些未知标签做出任何假设并且承认它们是未知的，在模型训练时的损失函数中最大化深度多标签分类网络对于未知标签预测概率的信息熵，用于解决 SPML 问题，使得模型可以免受假阴性标签噪声的影响，利用真实正标签提供的监督信息进行学习，并对未知标签做出可区分性较强的预测，EMR 方法易于实现，不会引入额外的可学习参数，同时能够与任意的深度多标签分类网络相结合，性能优异，适宜处理超大规模多标签图像数据集。

本发明相比于 ROLE 方法，ROLE 方法引入的 EPR 正则化方法需要提前统计数据集中每个样本的平均真实正标签数量，然而这在实际应用场景中难以实现，因为数据集中真实正标签的数目是未知的。此外，ROLE 方法采用的标签估计模块给原本的网络模型引入额外的大量的可学习参数，大大增加网络训练时的时间和空间需求，且引入的可学习参数数目与数据集中的类别数目和样本数目正相关，这表明 ROLE 方法不适用于超大规模多标签图像数据集。而本发明提出的方法不会引入额外的可学习参数，适用场景更加广泛，可以与任意的深度多标签分类网络和任意规模的多标签数据集结合使用，同时性能远远优于 ROLE 方法。

为了验证本发明的优越性，采用本发明EMR方法和现有技术的AN方法和ROLE方法分别在PASCAL VOC 2012 (VOC)、MS-COCO 2014 (COCO)、NUS-WIDE (NUS)和CUB-200-2011 (CUB)这四个大规模多标签图像数据集上进行实验验证，实验结果表明本发明所提出的方法性能优异，超越了所有的现存技术方案。实验结果为每个数据集中的最优性能，如下表所示：

数据集 / 方法	VOC	COCO	NUS	CUB
AN	85.89	64.92	42.27	18.31
ROLE	87.77	67.04	41.63	13.66
EMR	89.09	70.70	47.15	20.85

本发明旨在解决每个样本仅有单个正标签且其余标签均缺失的情况下，深度多标签分类网络的鲁棒学习问题。本发明提出了一种基于信息熵最大化的二元交叉熵损失函数正则化机制。利用本发明所提出的方法进行训练，深度多标签分类网络的泛化性能可以大幅提高，在极度标签缺失的情况下表现良好。

以上所述仅为本发明的优选实施例而已，并不用于限制本发明，对于本领域的技术人员来说，本发明可以有各种更改和变化。凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

权 利 要 求 书

1、基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，其特征在于，包括：深度多标签分类网络对于每一个样本的每一个未知标签的预测均视为一个离散随机变量，包括预测为正标签和预测为负标签两种事件，预测为正标签的概率即为深度多标签分类网络输出的未知标签预测概率，控制深度多标签分类网络输出的未知标签预测概率的信息熵最大化。

2、根据权利要求 1 所述的基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，其特征在于，所述方法中假设给定一个仅有单个正标签的样本对$(\mathbf{x}^{(n)}, \mathbf{z}^{(n)})$，其中$\mathbf{x}^{(n)}$为第$n$个输入图像，$\mathbf{z}^{(n)}$为第$n$个图像对应的标签向量，则损失函数为：

$$L_{\mathrm{EMR}}(\mathbf{f}^{(n)}, \mathbf{z}^{(n)}) = -\frac{1}{C}\sum_{c=1}^{C}[1_{[z_c^{(n)}=1]}\log(f_c^{(n)}) + 1_{[z_c^{(n)}=0]}R(f_c^{(n)})]$$

其中，$\mathbf{f}^{(n)}$是模型关于$\mathbf{x}^{(n)}$的输出，$f_c^{(n)}$是模型在第c个类上的预测概率，C是类别数目，$z_c^{(n)}=1$表示第c个类存在正标签，$z_c^{(n)}=0$表示第c个类存在负标签，$R(f_c^{(n)})$为正则化项。

3、根据权利要求 2 所述的基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，其特征在于，所述正则化项$R(f_c^{(n)})$定义为：

$$R(f_c^{(n)}) = -\alpha[f_c^{(n)}\log(f_c^{(n)}) + (1-f_c^{(n)})\log(1-f_c^{(n)})]$$

其中，α为控制正则化强度大小的超参数。

4、根据权利要求 3 所述的基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，其特征在于，所述超参数α的值在 0 到 1 之间进行调节。

5、根据权利要求 4 所述的基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，其特征在于，所述方法中预测为正标签的概率即为网络输出的未知标签预测概率$f_c^{(n)}$，预测为负标签的概率即为$1-f_c^{(n)}$。

6、根据权利要求 1 所述的基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，其特征在于，所述方法控制未知标签预测概率趋于 0.5。

7、根据权利要求 1 所述的基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，其特征在于，所述方法中假设有一个离散随机变量$X \in \{x_1, x_2, ..., x_n\}$，则其信息熵的定义为：

$$H(X) = \mathrm{E}[I(X)] = \mathrm{E}[-\ln(P(X))]$$

其中，$P(X)$为X的概率质量函数，$E[\cdot]$为期望函数，$I(X)$即为X的信息量。

8、根据权利要求7所述的基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，其特征在于，所述方法中若已知或已对$P(X)$做出估计后，则信息熵的公式表示为：

$$H(X)=\sum_i P(x_i)I(x_i)=-\sum_i P(x_i)\log_b P(x_i)$$

其中，b 是对数所使用的底。

9、根据权利要求 8 所述的基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，其特征在于，所述方法中当 $b=2$，信息熵的单位是 bit；当 $b=e$，信息熵的单位是 nat；当 $b=10$，信息熵的单位是 Hart。

10、根据权利要求 9 所述的基于信息熵最大化正则机制的深度多标签分类网络鲁棒训练方法，其特征在于，所述方法中深度多标签分类网络对于每一个样本的每一个未知标签的预测均视为一个离散随机变量 $X=\{x_1,x_2\}$，事件 x_1 即为预测为正标签，事件 x_2 即为预测为负标签，则 $P(x_1)$ 即为网络输出的未知标签预测概率，$P(x_2)=1-P(x_1)$ 即为预测为负标签的概率。

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2021/137348

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/62(2022.01)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; G06N

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)

WPI; EPODOC; CNKI; CNPAT: 熵, 最大, 正则, 标签, 随机, 变量, 损失, 函数, entropy, maximum, regular, label, random, variable, loss, function

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2020334565 A1 (SIEMENS AKTIENGESELLSCHAFT) 22 October 2020 (2020-10-22) description, paragraphs [0005]-[0040], and figures 1-8	1-10
A	US 2020160177 A1 (ROYAL BANK OF CANADA) 21 May 2020 (2020-05-21) entire document	1-10
A	CN 110163234 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 23 August 2019 (2019-08-23) entire document	1-10

☐ 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	Date of mailing of the international search report
01 September 2022	**14 September 2022**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2021/137348

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
US 2020334565 A1	22 October 2020	None	
US 2020160177 A1	21 May 2020	CA 3061717 A1	16 May 2020
CN 110163234 A	23 August 2019	WO 2020073951 A1	16 April 2020
		US 2021042580 A1	11 February 2021

国际检索报告

国际申请号
PCT/CN2021/137348

A. 主题的分类

G06K 9/62(2022.01)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; G06N

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）)

WPI;EPODOC;CNKI;CNPAT:熵，最大，正则，标签，随机，变量，损失，函数，entropy，maximum，regular，label，random，variable，loss，function

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
X	US 2020334565 A1 (SIEMENS AKTIENGESELLSCHAFT) 2020年10月22日 (2020-10-22) 说明书第[0005]-[0040]段，附图1-8	1-10
A	US 2020160177 A1 (ROYAL BANK OF CANADA) 2020年5月21日 (2020-05-21) 全文	1-10
A	CN 110163234 A (腾讯科技深圳有限公司) 2019年8月23日 (2019-08-23) 全文	1-10

☐ 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	国际检索报告邮寄日期
2022年9月1日	2022年9月14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 冯婷霆 电话号码 86-(10)-53961364

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2021/137348

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)
US 2020334565 A1	2020年10月22日	无	
US 2020160177 A1	2020年5月21日	CA 3061717 A1	2020年5月16日
CN 110163234 A	2019年8月23日	WO 2020073951 A1	2020年4月16日
		US 2021042580 A1	2021年2月11日