

一箭N雕：多任务深度学习实战



程程 · 1 年前

深度学习大讲堂致力于推送人工智能，深度学习方面的最新技术，产品以及活动。请关注我们



1、多任务学习导引

多任务学习是机器学习中的一个分支，按1997年综述论文Multi-task Learning一文的定义：Multitask Learning (MTL) is an inductive transfer mechanism whose principle goal is to improve generalization performance. MTL improves generalization by leveraging the domain-specific information contained in the training signals of related tasks. It does this by training tasks in parallel while using a shared representation。翻译成中文：**多任务学习是一种归纳迁移机制，基本目标是提高泛化性能。多任务学习通过相关任务训练信号中的领域特定信息来提高泛化能力，利用共享表示采用并行训练的方法学习多个任务。**

顾名思义，多任务学习是一种同时学习多个任务的机器学习方法，如图1所示，多任务学习同时学习了人类和狗的分类器以及男性和女性的性别分类器。

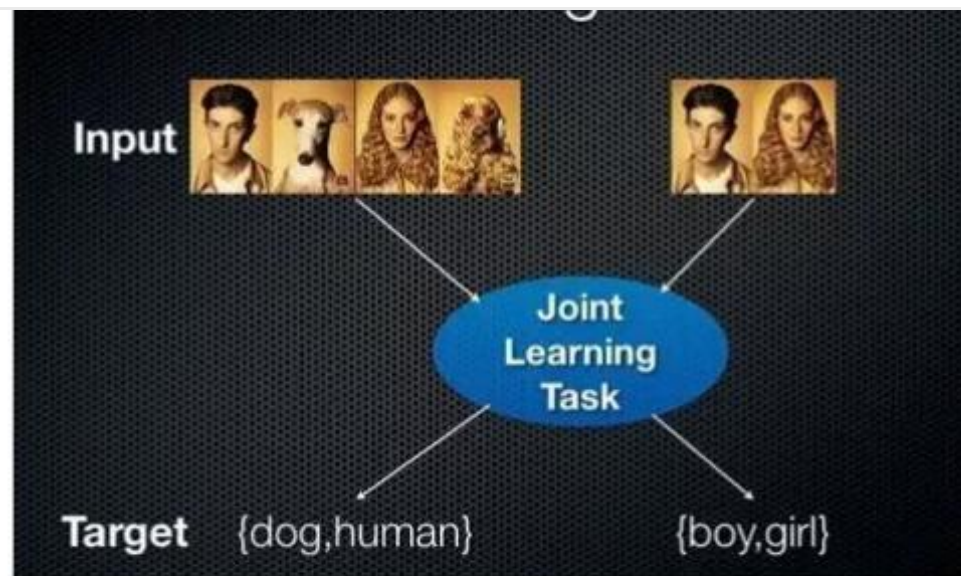


图1. 多任务学习示例

(图片引自：<http://www.cs.cornell.edu/~kilian/research/multitasklearning/multitasklearning.html>)

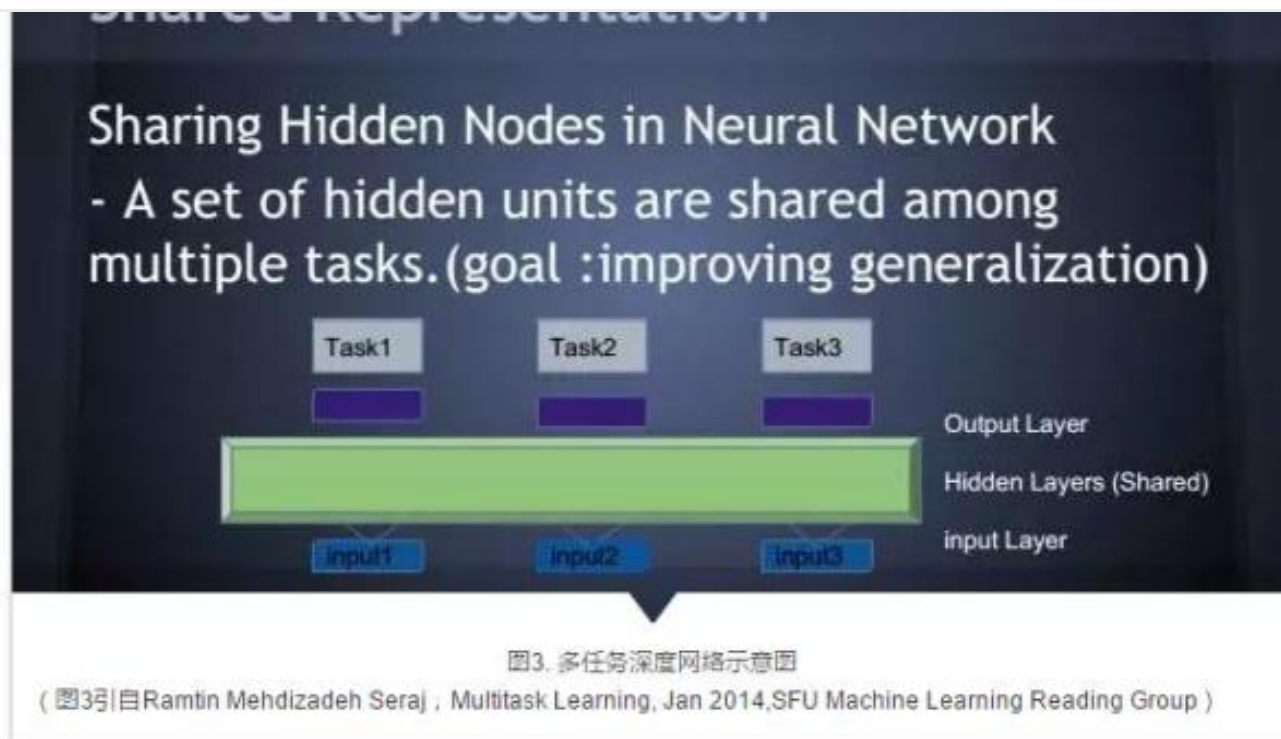
进一步的，图2所示为单任务学习和多任务学习的对比。在单任务学习中，每个任务采用单独的数据源，分别学习每个任务单独的模型。而多任务学习中，多个数据源采用共享表示同时学习多个子任务模型。



多任务学习的基本假设是多个任务之间具有相关性，因此能够利用任务之间的相关性互相促进。例如，属性分类中，抹口红和戴耳环有一定的相关性，单独训练的时候是无法利用这些信息，多任务学习则可以利用任务相关性联合提高多个属性分类的精度，详情可参考文章Maryland大学Hand等人的论文Attributes for Improved Attributes: A Multi-Task Network for Attribute Classification。

2、多任务深度学习

近年来，在深度学习技术的推动下计算机视觉领域取得了突飞猛进的进展。本质上说，深度学习是多层的神经网络，对输入进行了层级的非线性表示，来自网络可视化的证据表明，深度网络的层级表示从语义上从底层到高层不断递进。深度网络强大的表示能力，使得多任务深度学习有了施展的空间。图3所示为多任务深度网络结构示意图。Input x表示不同任务的输入数据，绿色部分表示不同任务之间共享的层，紫色表示每个任务特定的层，Task x表示不同任务对应的损失函数层。在多任务深度网络中，低层次语义信息的共享有助于减少计算量，同时共享表示层可以使得几个有共性的任务更好的结合相关性信息，任务特定层则可以单独建模任务特定的信息，实现共享信息和任务特定信息的统一。



在深度网络中，多任务的语义信息还可以从不同的层次输出，例如GoogLeNet中的两个辅助损失层。另外一个例子比如衣服图像检索系统，颜色这类的信息可以从较浅层的时候就进行输出判断，而衣服的风格样式这类的信息，更接近高层语义，需要从更高的层次进行输出，这里的输出指的是每个任务对应的损失层的前一层。

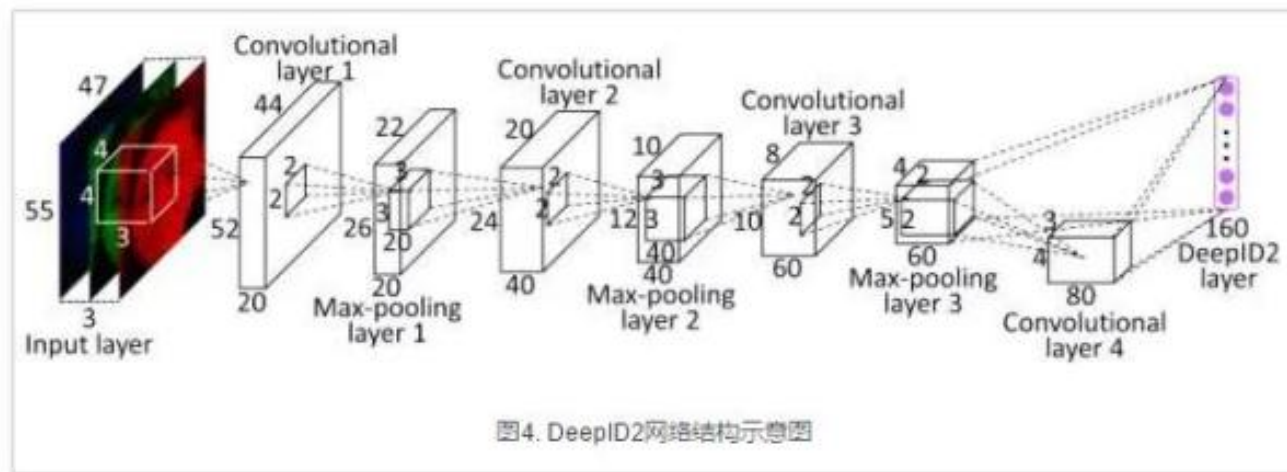
3、多任务深度学习应用案例

目前，多任务深度学习已经广泛应用于人脸识别、细粒度车辆分类、面部关键点定位与属性分类等多个领域，以下介绍其中的代表性论文。

3.1 人脸识别网络 DeepID2



别网络DeepID2，网络结构如下图所示：



DeepID2中共有两个损失函数，分别为人脸分类损失函数，对应于Caffe中的SoftmaxLoss：

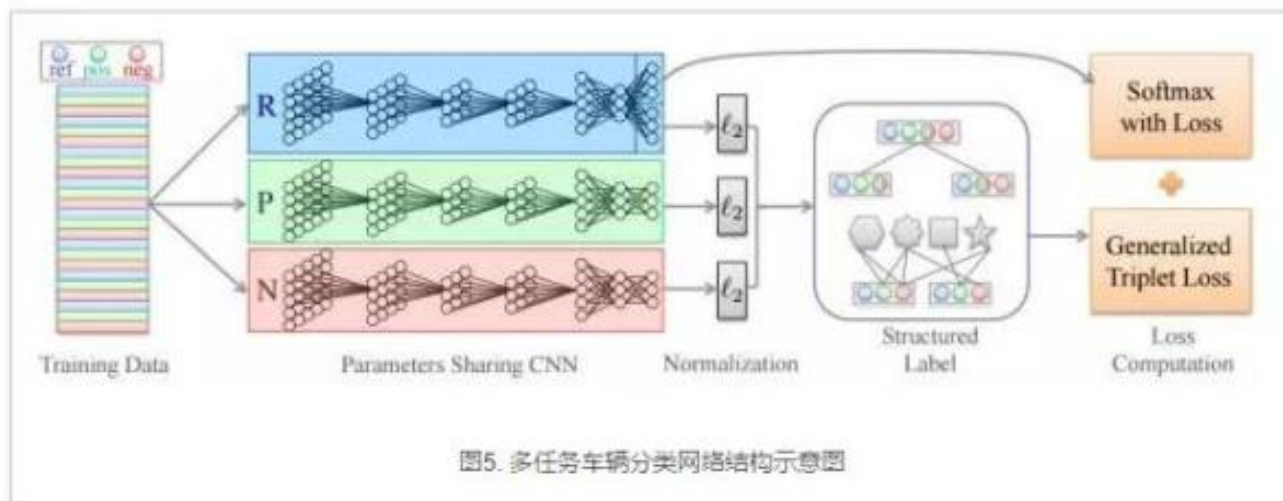
$$\text{Ident}(f, t, \theta_{id}) = - \sum_{i=1}^n -p_i \log \hat{p}_i = - \log \hat{p}_t$$

另外一个为是人脸确认损失函数，对应于Caffe中的Contrastive Loss：

$$\text{Verif}(f_i, f_j, y_{ij}, \theta_{ve}) = \begin{cases} \frac{1}{2} \|f_i - f_j\|_2^2 & \text{if } y_{ij} = 1 \\ \frac{1}{2} \max(0, m - \|f_i - f_j\|_2)^2 & \text{if } y_{ij} = -1 \end{cases}$$

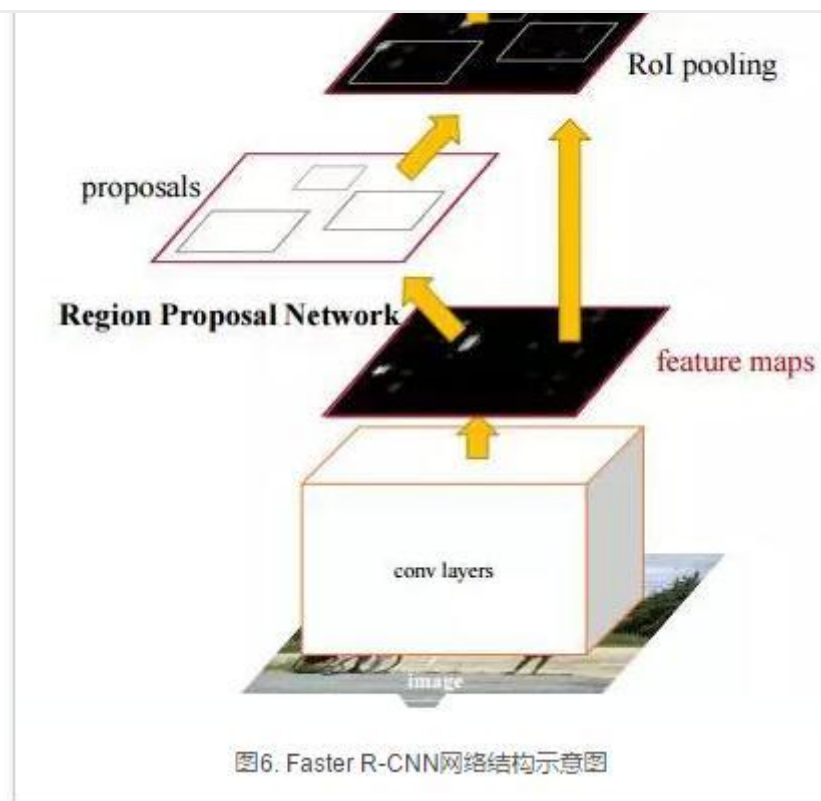
3.2细粒度车辆分类网络

arXiv。作者将这个网络用于细粒度车辆分类上，提醒注意的是为了计算Triplet Loss，特征进行了L2范数归一操作，网络结构如下图所示：



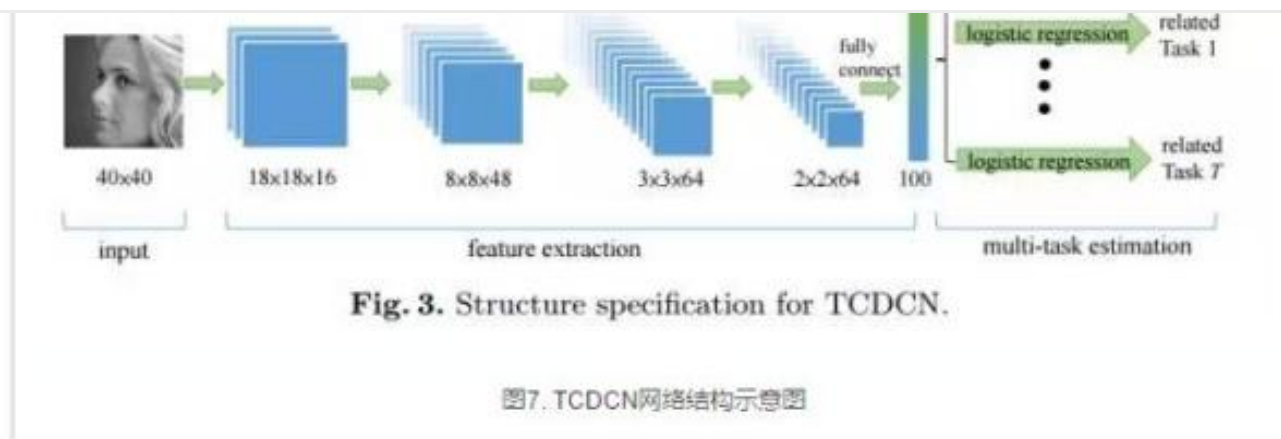
3.3物体检测网络Faster R-CNN

在物体检测网络Faster R-CNN中也有多任务学习的应用。Faster R-CNN的网络结构如下图6所示，包含两个任务，分别为窗口回归和窗口分类，其中RPN模块的卷积层在两个任务之间共享。Faster R-CNN的最新版本支持整体端到端训练，可以同时检测多类物体，是目前最具代表性的目标检测框架，同时也是多任务深度学习的一个典型应用。



3.4面部关键点定位与属性分类网络TDCN

面部关键点估计和头部姿态以及人脸属性（是否戴眼镜、是否微笑和性别）之间有着紧密的联系，香港中文大学汤晓鸥组发表于ECCV14的工作Facial Landmark Detection by Deep Multi-task Learning利用多任务学习方法联合进行人脸面部关键点定位和属性预测，网络结构如下图7所示。



4、基于Caffe实现多任务学习的小样例

本节在目前广泛使用的深度学习开源框架Caffe的基础上实现多任务深度学习算法所需的多维标签输入。默认的，Caffe中的Data层只支持单维标签，为了支持多维标签，首先修改Caffe中的convert_imageset.cpp以支持多标签：



```
std::string filename;
std::string label_count_string = argv[5];
int label_count = std::atoi(label_count_string.c_str());
std::vector<float> label(label_count);
while (infile >> filename)
{
    for (int i = 0; i < label_count; i++)
        infile >> label[i];
    lines.push_back(std::make_pair(filename, label));
}
// Create new DB
scoped_ptr<db::DB> db_image(db::GetDB(FLAGS_backend));
scoped_ptr<db::DB> db_label(db::GetDB(FLAGS_backend));
db_image->Open(argv[3], db::NEW);
db_label->Open(argv[4], db::NEW);
scoped_ptr<db::Transaction> txn_image(db_image->NewTransaction());
scoped_ptr<db::Transaction> txn_label(db_label->NewTransaction());
```

这样我们就有了多任务的深度学习的基础部分数据输入。为了向上兼容Caffe框架，本文摒弃了部分开源实现增加Data层标签维度选项并修改Data层代码的做法，直接使用两个Data层将数据读入，即分别读入数据和多维标签，接下来介绍对应的网络结构文件prototxt的修改，注意红色的注释部分。



```
name: "data"
type: "Data"
top: "data" #注意这里，top层输出blob只剩一个
transform_param {
  mean_file: "/home/xyf/Multilabel/mean.binaryproto"
  crop_size: 227
}
include {
  phase: TRAIN
}
data_param {
  source: "/home/xyf/Multilabel/TrainDbImage"
  batch_size: 128
  backend: LMDB
}
}
#这是我们使用的标签数据库，用来负责存放标签
layer {
  name: "labels"
  type: "Data"
  top: "labels"
  include {
    phase: TRAIN
  }
  data_param {
    source: "/home/xyf/Multilabel/TrainDbLabel"
    batch_size: 128
    backend: LMDB
  }
}
layer {
  name: "data"
  type: "Data"
  top: "data"
  transform_param {
    mean_file: "/home/xyf/Multilabel/mean.binaryproto"
```



```
phase: TEST
}
data_param {
  source: "/home/xyf/Multilabel/TestDbImage"
  batch_size: 128
  backend: LMDB
}
}
layer {
  name: "labels"
  type: "Data"
  top: "labels"
  include {
    phase: TEST
  }
  data_param {
    source: "/home/xyf/Multilabel/TestDbLabel"
    batch_size: 128
    backend: LMDB
  }
}
```

特别的，slice层对多维的标签进行了切分，为每个任务输出了单独的标签。



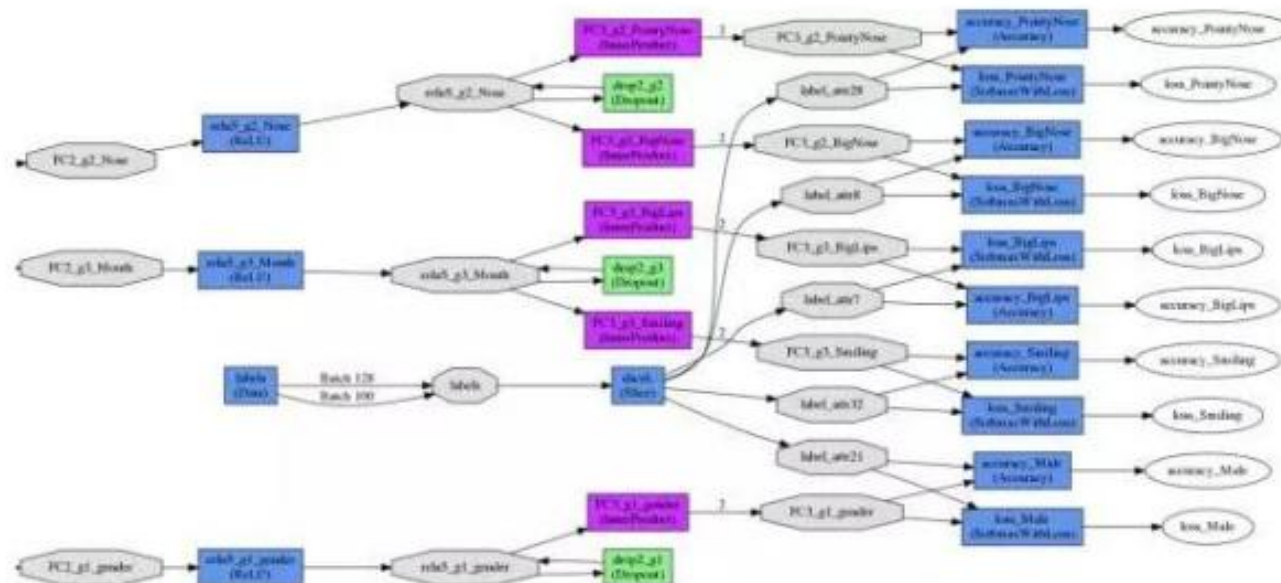
```
name: slice
type: "Slice"
bottom: "labels"
top: "label_attr1"
top: "label_attr2"
top: "label_attr3"
top: "label_attr4"
top: "label_attr5"
slice_param
{
  slice_dim: 1
  slice_point: 1
  slice_point: 2
  slice_point: 3
  slice_point: 4
}
}
```

另外一个值得讨论的是每个任务的权重设置，在本文实践中五个任务设置为等权重 loss_weight:0.2。一般的，建议所有任务的权重值相加为1，如果这个数值不设置，可能会导致网络收敛不稳定，这是因为多任务学习中对不同任务的梯度进行累加，导致梯度过大，甚至可能引发参数溢出错误导致网络训练失败。

```
#且看loss_weight，本文作者选择5个属性的重要性一样，所以每个0.2。
layer {
  name: "loss_attr1"
  type: "SoftmaxWithLoss"
  bottom: "bottom_predict1"
  bottom: "label_attr1"
  top: "loss_attr1"
  loss_weight: 0.2
}
```


[CodeSnap/convert_multilabel.cpp at master · HolidayXue/CodeSnap · GitHub](#)

多任务损失函数层的网络结构示意图如下图所示：



5. 总结

本文回顾了多任务学习的基本概念，并讨论了多任务深度学习的基本思想和应用案例。最后以开源深度学习平台Caffe为例讨论了多任务深度学习的实现，并给出了开源代码。

致谢

本文在投稿之后经历了三轮修改，其中一轮公众号编辑部初审，一轮双盲评审大改和一轮单盲评审小修，两名审稿专家对原文进行了全面仔细的阅读，帮助作者修正了文章的若干理论表



该文章属于“深度学习大讲堂”原创，如需要转载，请联系@果果是枚开心果.

作者简介：



薛云峰，([HolidayXue \(HolidayXue\) · GitHub](#))，主要从事视频图像算法的研究，就职于浙江捷尚视觉科技股份有限公司担任深度学习算法研究员。捷尚致力于视频大数据和视频监控智能化，现诚招业内算法和工程技术人才，招聘主页[浙江捷尚视觉科技股份有限公司--安全服务运营商](#)，联系邮箱：hr@icarevision.cn

原文链接：[一箭N雕：多任务深度学习实战](#)

知



首发于
深度学习大讲堂

写文章

登录



欢迎关注我们深度学习大讲堂的微信公众号。

深度学习（Deep Learning）

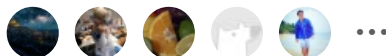


39

☆ 收藏

分享

⚠ 举报



文章被以下专栏收录



深度学习大讲堂

人工智能的产品技术前沿

进入专栏

2 条评论

写下你的评论...



李作新

多任务损失函数层的网络结构示意图分辨率太低。。



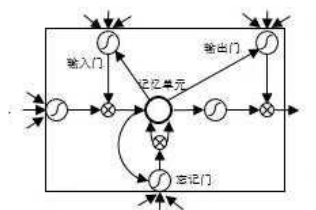
猿族 回复 李作新

同感

1 年前

查看对话

推荐阅读



(b) LSTM 网络[Hochreiter97]
典型的递归神经网络模型

【青年学者专栏】递归神经网络(Recurrent Neural Network)学习

深度学习大讲堂致力于推送人工智能，深度学习方面的最新技术，产品以及活动。请关注我们的知... [查看全文](#) >

程程 · 1 年前 · 发表于 深度学习大讲堂



天津大学深度学习一线实战研讨班干货总结与资源下载

深度学习大讲堂致力于推送人工智能，深度学习方面的最新技术，产品以及活动。请关注我们的知... [查看全文](#) >

程程 · 1 年前 · 发表于 深度学习大讲堂

知



首发于
深度学习大讲堂

写文章

登录



同日，上海一中法院对于一起涉及星娱乐法公司的案件，判决书中

同尚存在履约可能，因此判决... 查看全文 >

李振武 · 3 个月前 · 编辑精选 · 发表于 星娱乐法



混动五讲之三：混动神器——行星齿轮组

本文首发于微信公众号『驾仕派』，转载请与我联系系列文章：混动五讲之一：混动到底靠什么省... 查看全文 >

鱼非鱼 · 20 天前 · 编辑精选