

# Constrained Neural Network and Its Application in Target Direction Estimation

Jin Liu, Shumin Liu\*, Gao Fang

State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan Hubei

Email: [liushumin@whu.edu.cn](mailto:liushumin@whu.edu.cn)

Received: Jun. 11<sup>th</sup>, 2018; accepted: Jun. 28<sup>th</sup>, 2018; published: Jul. 5<sup>th</sup>, 2018

---

## Abstract

At present, the mainstream convolutional neural network (CNN) approaches focus on the recognition and positioning of targets in the field of computer vision. Most locate the position of targets with axis-aligned bounding boxes. In order to make a deeper understanding of the images and obtain more accurate position of various targets, the direction information is needed. Therefore, a new method is proposed to estimate the target direction, which applies CNN to regress the two directional components of the target direction angle that describes the direction, instead of directly regressing the direction angle like some existing approaches which shows some shortcomings. Considering the function constraint of the square sum of 1 between the directional components, this paper proposes a general method to solve this kind of problems with output constraints by using constrained neural network, which introduces constraint errors into the Loss layer and the back propagation process, and applies it in the target direction estimation. Experiments show that, the target direction estimation method maintains the descent rate and range of the original output loss and moreover, reduce the constraint error and improve the prediction accuracy. The significance lies in proposing a general method to solve a kind of problems with output constraints, which reduces constraint errors, and shows general applicability.

## Keywords

CNN, Output Constraints, Target Direction Estimation, Directional Components

---

# 约束性神经网络及其在目标方向估计中的应用

刘 进, 刘淑敏\*, 方 高

武汉大学测绘遥感信息工程国家重点实验室, 湖北 武汉

Email: [liushumin@whu.edu.cn](mailto:liushumin@whu.edu.cn)

收稿日期: 2018年6月11日; 录用日期: 2018年6月28日; 发布日期: 2018年7月5日

\*通讯作者。

## 摘要

目前,在计算机视觉领域,主流的卷积神经网络算法专注于目标的识别和定位,且大多数采用轴对齐包围盒定位目标,而为了对图像做更深刻的语义理解,更加精准地定位各类目标,需要获取目标的方向信息。因此,本文提出一种针对图像目标方向估计的方法,采用卷积神经网络对描述目标方向的两个方向分量进行回归,规避一些现有方向估计方法直接对方向角回归而产生的缺点。由于方向分量之间存在平方和为1的函数约束,本文提出约束性神经网络的概念,进一步提出利用约束性神经网络解决这类带有输出约束问题的一般性方法,即在Loss层引入约束误差,参与反向传播,并将其具体运用于目标方向估计中。经实验,本文采用的基于约束性神经网络的目标方向估计方法,能够在保证原输出损失的下降速度和幅度的前提下,降低约束误差,提高估计精度。

## 关键词

卷积神经网络, 输出约束, 目标方向估计, 方向分量

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

近年来,神经网络发展迅速,尤其是卷积神经网络(Convolutional Neural Network, CNN) [1],已经广泛应用于计算机视觉、语音识别、医学研究以及智能博弈等诸多领域,且仍在不断更新迭代。对于其中一个热点问题——图像目标检测,R-CNN [2]是将CNN方法引入进来的一个里程碑,随后衍生出系列研究Fast R-CNN [3]、Faster R-CNN [4]等,2016年YOLO [5]的提出,将目标的分类和位置检测作为一个整体回归问题进行求解,在识别速度上实现了重大突破,接着,SSD [6]、YOLO9000 [7]等进一步提高了目标检测的速度和精度。

然而人们并不仅仅满足于获取目标的类别和位置信息,还希望从图像中挖掘更多的信息,例如目标的方向信息,这样就可以对图像做出更深刻的语义理解,产生更多的应用,例如利用目标的方向信息对图像进行纠正,预测与目标特定方向相关联事件的发生概率,再如估计出动态目标的方向信息就能预测其意图等。另外,对于在原始图像上标记目标所在位置,上述方法均采用传统的轴对齐包围盒来标记目标,而针对任意角度倾斜的目标,特别是倾斜的细长条状物体,如船只、车辆、铅笔等,如果可以预测出目标的方向,更适合使用带有方向信息的倾斜包围盒进行定位,这样可以更加准确地包含目标,减少框内背景信息比例,增加包围盒的信息利用率,同时在目标密集区域也可大大减少多个包围盒之间的重叠率,更有效地区分不同目标个体,传统包围盒与倾斜包围盒的对比图如图1所示。

所以,对目标方向估计的研究很有价值。目前,针对目标方向估计也有不少研究,如文献[8]提出 $R^2$ CNN检测任意方向的文字目标,先粗预测文字的轴对齐包围盒,再通过不同尺度池化层提取的特征,精化轴对齐包围盒,再得到最小倾斜包围盒,非直接预测倾斜包围盒;文献[9]采用有方向的SSD检测任意方向的车辆,直接将描述方向的方向角纳入到目标包围盒的属性数据中进行回归;文献[10]则预先设定几组特定角度的旋转盒,预测时在各个位置基于这些预设的旋转盒及不同类型目标比例特征分子任务进行预测,且包围盒中也是添加一项方向角描述方向。但是,由于方向角度值 $\theta$ 在旋转一整周后,即从 $360^\circ$

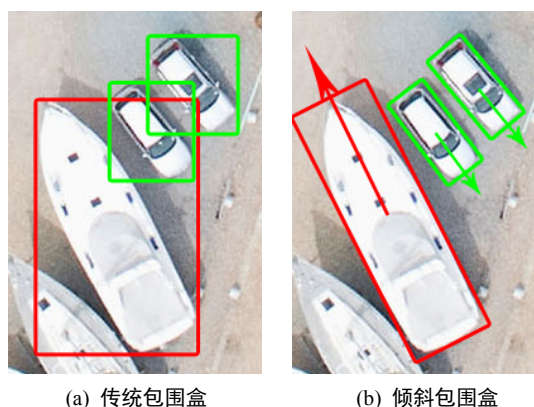


Figure 1. Comparison of traditional bounding box and rotated bounding box  
图 1. 传统包围盒与倾斜包围盒的对比图

到  $0^\circ$  处角度的数值存在突变，而实际上  $360^\circ$  和  $0^\circ$  是指的同一方向，所以如果直接使用  $\theta$  进行回归，网络在  $0^\circ$  和  $360^\circ$  处的预测值在数值上极有可能会存在很大误差。

本文采用目标方向  $\theta$  的两个单位方向分量  $(x, y)$  ( $\sin\theta, \cos\theta$ )，作为卷积神经网络的输入数据和预测数据，因为  $\sin\theta$  和  $\cos\theta$  在  $\theta \in [0^\circ, 360^\circ)$  上具有单调性和连续性，不会出现上述突变问题。而由于方向分量为单位矢量，即两个输出分量存在平方和为 1 的函数约束，这就会引入一个新问题，即如何确保神经网络输出的矢量是单位矢量。传统神经网络的主体思想是建立从输入到输出的数学模型，通过大量有标注的样本数据对模型进行训练，解算出其中未知的模型参数，最终得到完整的模型。而在诸多实际应用中，网络输出往往存在很多约束，例如需要满足某一值域范围，如图像目标框的长宽为正数，或是需要满足某种函数关系式，比如本文方向估计方法中的输出矢量分量平方和等于 1，本文将这种具有输出约束的神经网络称作约束性神经网络，探讨使用约束性神经网络解决一类具有约束问题的一般性方法，将其具体应用于目标方向估计中。

## 2. 约束性神经网络

传统神经网络的主体思想是，给定一个从输入  $\mathbf{X}$  到输出  $\mathbf{Y}$  的数学模型  $\mathbf{Y} = f(\mathbf{W}, \mathbf{X})$ ，其中  $\mathbf{X}$  是输入矢量， $\mathbf{Y}$  是输出矢量， $\mathbf{W}$  是参数矢量，通过给定大量样本数据集  $(x_i, y_i)$ ，其中  $i = 1, 2, \dots, n$ ，列出联合方程式，使用优化算法解算出模型参数  $\mathbf{W}$ ，最终建立从输入  $\mathbf{X}$  到输出  $\mathbf{Y}$  的完整数学模型。

然而，在不同的实际问题中，输入  $\mathbf{X}$  与输出  $\mathbf{Y}$  之间往往关系复杂，并非简单的线性关系可以表示，且输出  $\mathbf{Y}$  多会有值域上的范围约束(本文中称作弱约束)，或多个输出分量之间存在函数约束(本文中称作强约束)，所以不论用多少样本都是有限的，无法完整和精确地描述模型，对于一个新样本  $x_i^*$ ，经过神经网络前向传播输出的  $y_i^* = f(\mathbf{W}, x_i^*)$ ，不可能永远满足各项约束，这就提出一类新的问题——如何对现有神经网络进行改进，使其不仅符合有限样本集的约束，而且满足各种实际问题的约束，进而使得一些以前不方便直接用神经网络解决的问题也能得到解决。因此本文的意义在于提出一种解决一类具有约束问题的一般性方法。

### 2.1. 弱约束

弱约束是指输出  $\mathbf{Y}$  在值域上存在范围约束。例如，预测图像目标定位框的宽高，输出的值域范围是  $[0, +\infty)$ ；利用人脸图像预测年龄，输出的值域范围可以设为  $[0, 100]$  等。

神经网络原始输出的值域范围是  $(-\infty, +\infty)$ ，为了满足弱约束，一般性解决方案是采用输出转换函数

$\phi(\mathbf{y})$ ，将原始输出变换到要求的值域区间。输出转换函数  $\phi(\mathbf{y})$  必须满足以下两个特性：

- 1) 单调，这样转换才具有唯一性；
- 2) 输出值域与问题期待值域范围一致。

在网络进行前向传播时，使用输出转换函数对原始输出进行转换，对误差进行反向传播时根据链式求导法则求出对应的偏导值，进行权值修正。

事实上，常用的几个激励函数可以作为输出转换函数使用，对于值域为  $(-\infty, +\infty)$  的原始输出， $\tanh$  函数可将其映射到  $[-1, 1]$  区间上； $\text{sigmoid}$  函数可将其映射到  $[0, 1]$  区间上，而为了避免梯度消失的问题，现在很少在中间层使用； $\text{ReLU}$  函数可将正数保持不变，非正数全部映射成 0，由于会丢失掉负数部分的信息，不宜作为输出转换函数。另外，也可根据实际问题，构造合适的输出转换函数，例如在预测图像目标定位框的宽高问题上，可在前向传播时，使用指数函数  $\phi(y_i) = e^{y_i}$  对神经网络输出值进行转换，将  $y_i \in (-\infty, +\infty)$  映射到  $\phi(y_i) \in (0, +\infty)$  上。

## 2.2. 强约束

强约束是指根据问题需要，对输出矢量  $\mathbf{Y}$  的内部某些分量有函数约束，以  $\varphi(\mathbf{y}) = 0$  表示。对于不同问题，约束函数  $\varphi(\mathbf{y})$  的形式各不相同。

如果约束函数关系简单，比如只存在线性关系，也可通过输出转换函数实现。在经典的图像目标分类问题中，需要输出目标属于各个类别的概率，且概率总和为 1，取其中概率最大者对应的类别作为该目标的预测类别，此问题中输出  $\mathbf{Y}$  的各个分量需满足  $\sum_{i=1}^n y_i = 1$ ， $y_i \in [0, 1]$ ，即输出约束函数  $\varphi(\mathbf{y}) = \sum_{i=1}^n y_i - 1 = 0$ ， $y_i \in [0, 1]$ ，由于

$$\text{softmax}(y_i) = \frac{e^{y_i}}{\sum_{j=1}^n e^{y_j}} \quad (1)$$

能很好地满足该约束要求，所以使用  $\text{softmax}$  函数作为转换输出函数，在神经网络最后加一个  $\text{softmax}$  层即可解决该分类问题。

但是，在很多复杂情况下，约束函数非线性，很难找到合适的输出转换函数，由此，本文提出一种解决这类问题的一般性方案。

当输出矢量  $\mathbf{Y}$  的分量间有单个强约束函数关系  $\varphi(\mathbf{y}) = 0$  时，那么在构建神经网络时，不仅需要最终预测的  $\tilde{y}_i$  分量与真值  $\bar{y}_i$  作比较求误差，以  $\text{Diff}_y$  表示，则

$$\text{Diff}_y = \tilde{y}_i - \bar{y}_i \quad (2)$$

还需要考虑  $\tilde{y}_i$  分量之间的约束关系，增加一个约束误差计算，即计算  $\varphi(\tilde{\mathbf{y}})$  并与约束值真值 0 相减求误差，定义约束误差函数为  $\text{Error}_\varphi$ ，则

$$\text{Error}_\varphi = \varphi(\tilde{\mathbf{y}}) - 0 = \varphi(\tilde{\mathbf{y}}) \quad (3)$$

考虑约束误差的完整网络结构如图 2 所示。

在利用大量样本集进行网络训练的过程中，通过误差反向传播对各权值进行修正，反向传播采用经典的梯度下降法[11]，计算式如下。

使用均方差定义输出损失函数  $\text{Loss}_y$ ，则

$$\text{Loss}_y = \frac{1}{2} \sum_{i=1}^n (\tilde{y}_i - \bar{y}_i)^2 \quad (4)$$

$\text{Loss}_y$  在反向传播时，对  $\tilde{y}_i$  的偏导为

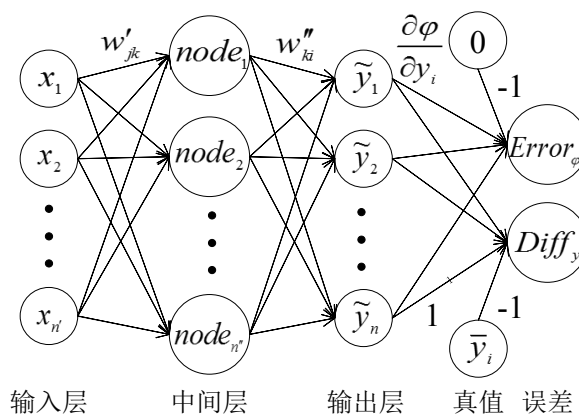


Figure 2. Neural network structure with single output constraint  
图 2. 带有单个输出约束的神经网络结构图

$$\frac{\partial Loss_y}{\partial \tilde{y}_i} = \tilde{y}_i - \bar{y}_i \quad (5)$$

约束误差的平方为

$$Error_\phi^2 = (\phi(\tilde{\mathbf{y}}) - 0)^2 = \phi^2(\tilde{\mathbf{y}}) \quad (6)$$

$Error_\phi^2$  在反向传播时, 对  $\tilde{y}_i$  的偏导为

$$\frac{\partial Error_\phi^2}{\partial \tilde{y}_i} = 2\phi(\tilde{\mathbf{y}}) \times \frac{\partial \phi(\tilde{\mathbf{y}})}{\partial \tilde{y}_i} \quad (7)$$

记总体误差为  $E_{\text{sum}}$ , 则

$$E_{\text{sum}} = \lambda \cdot Error_\phi^2 + Loss_y \quad (8)$$

其中  $\lambda$  为约束误差的常数系数。

则  $E_{\text{sum}}$  在反向传播时, 对  $\tilde{y}_i$  的偏导为

$$\frac{\partial E_{\text{sum}}}{\partial \tilde{y}_i} = \lambda \frac{\partial Error_\phi^2}{\partial \tilde{y}_i} + \frac{\partial Loss_y}{\partial \tilde{y}_i} \quad (9)$$

代入公式(5)和(7), 则  $E_{\text{sum}}$  对  $\tilde{y}_i$  的偏导为

$$\frac{\partial E_{\text{sum}}}{\partial \tilde{y}_i} = 2\lambda \cdot \phi(\tilde{\mathbf{y}}) \times \frac{\partial \phi(\tilde{\mathbf{y}})}{\partial \tilde{y}_i} + (\tilde{y}_i - \bar{y}_i) \quad (10)$$

权值修正根据导数的链式法则计算, 例如, 可通过下式更新权值  $w''_{ki}$ ,

$$\frac{\partial E_{\text{sum}}}{\partial w''_{ki}} = \frac{\partial E_{\text{sum}}}{\partial \tilde{y}_i} \times \frac{\partial \tilde{y}_i}{\partial w''_{ki}} \quad (11)$$

$$w''_{ki}{}^{+} = w''_{ki} - \eta \cdot \frac{\partial E_{\text{sum}}}{\partial w''_{ki}} \quad (12)$$

其中  $\eta$  为神经网络的学习率。

其他部分与传统神经网络处理方法一致, 由此对网络中所有权值进行更新迭代, 得到模型最佳参数矢量  $\mathbf{W}$ 。

考虑约束误差的反向传播过程如图 3 所示。

如果输出分量  $y_i$  之间具有  $N$  个输出约束  $\varphi_j(\mathbf{y})=0$ ，其中  $j=1,2,\dots,N$ ，同理，在网络中对应增加  $N$  个约束误差计算，每个约束误差函数为  $Error_{\varphi_j}$ ，则

$$Error_{\varphi_j} = \varphi_j(\tilde{\mathbf{y}}) - 0 = \varphi_j(\tilde{\mathbf{y}}) \tag{13}$$

构建如图 4 所示的带有  $N$  个输出约束的神经网络。

此时，总体误差为

$$E_{\text{sum}} = \sum_{j=1}^N \lambda_j \cdot Error_{\varphi_j}^2 + Loss_y \tag{14}$$

其中  $\lambda_j$  为各个约束误差的常数系数。

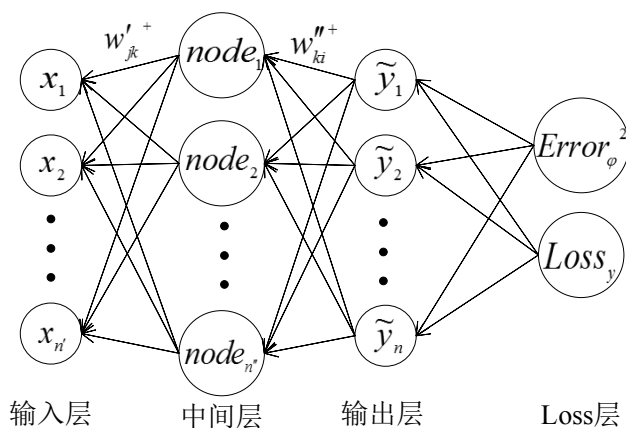


Figure 3. Back propagation of the neural network with single output constraint  
图 3. 带有单个输出约束的神经网络反向传播示意图

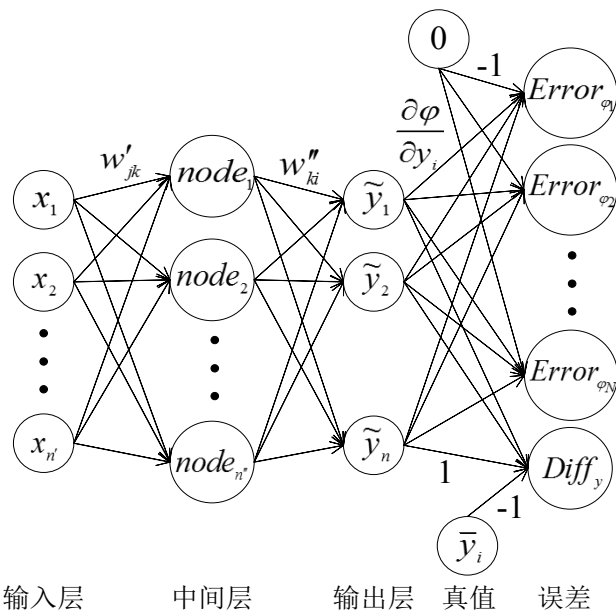


Figure 4. Neural network structure with multiple output constraints  
图 4. 带有  $N$  个输出约束的神经网络结构图

则  $E_{\text{sum}}$  在反向传播时, 对  $\tilde{y}_i$  的偏导为

$$\frac{\partial E_{\text{sum}}}{\partial \tilde{y}_i} = \sum_{j=1}^N \lambda_j \frac{\partial \text{Error}_{\phi_i}^2}{\partial \tilde{y}_i} + \frac{\partial \text{Loss}_y}{\partial \tilde{y}_i} = 2 \sum_{j=1}^N \lambda_j \cdot \phi_j(\tilde{\mathbf{y}}) \times \frac{\partial \phi_j(\tilde{\mathbf{y}})}{\partial \tilde{y}_i} + (\tilde{y}_i - \bar{y}_i) \quad (15)$$

同样通过链导原则进行权值修正和迭代更新, 最终构建完整模型。

以上即为输出带有一个或多个强约束关系的神经网络的一般处理方案。约束性问题模型广泛存在, 例如第二节将介绍的使用约束性神经网络预测目标方向分量, 进而估计目标方向。

### 3. 目标方向估计

#### 3.1. 网络输出设计

利用神经网络进行目标方向估计的主体思想是将目标方向作为卷积神经网络的一个输入对其进行训练, 最终进行预测。首先, 如图 5 所示, 本文中定义的目标方向指, 以朝上为标准方向, 即  $0^\circ$  方向, 顺时针旋转  $\theta$  角度为  $\theta$  方向,  $\theta \in [0^\circ, 360^\circ)$ 。

但是, 由于角度  $\theta$  在旋转一整周后, 即从  $360^\circ$  到  $0^\circ$  处角度的数值存在突变, 而实际上  $360^\circ$  和  $0^\circ$  是指的同一方向, 所以如果直接使用  $\theta$  角度值进行回归, 网络在  $0^\circ$  和  $360^\circ$  处的预测值在数值上极有可能会存在很大误差。所以本文采用目标方向  $\theta$  的两个单位方向分量  $(x, y)$  ( $\sin\theta, \cos\theta$ ), 作为卷积神经网络的输入数据和预测数据, 因为  $\sin\theta$  和  $\cos\theta$  在  $\theta \in [0^\circ, 360^\circ)$  上具有单调性和连续性, 不会出现上述问题, 并且用预测出的一组方向分量, 可以通过反三角函数公式并结合符号, 唯一确认目标方向  $\theta$ , 实现目标方向估计。

#### 3.2. 方向分量间的强约束

令  $x = \sin\theta$ ,  $y = \cos\theta$  作为网络的输出, 如图 5 所示, 由于方向分量为单位矢量, 即两个输出分量之间存在函数约束  $x^2 + y^2 = 1$ , 则目标方向估计神经网络就是输出带有单个强约束关系的神经网络, 其强约束函数  $\varphi$  为

$$\varphi = x^2 + y^2 - 1 = 0 \quad (16)$$

#### 3.3. 目标方向估计神经网络

为进行方向分量估计, 现建立一个从输入图像  $I$  到输出方向分量  $(x, y)$  的卷积神经网络数学模型

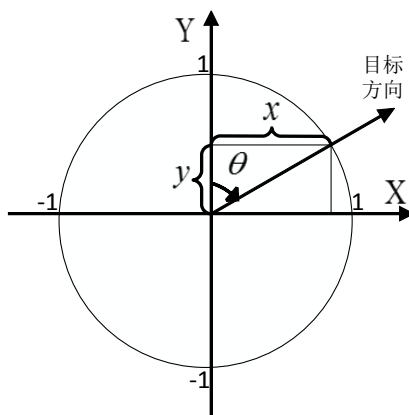


Figure 5. Direction angle  $\theta$  of the target

图 5. 目标方向角  $\theta$  示意图

$(x, y) = f(W, I)$ , 其中  $W$  是参数矢量。通过给定大量样本图片  $I_i$  和对应方向分量  $(x_i, y_i)$ , 其中  $i = 1, 2, \dots, n$ , 利用神经网络解算出模型参数  $W$ , 最终建立从输入图像  $I$  到输出方向分量  $(x, y)$  的完整映射, 实现对任一图像的方向估计。

本文神经网络结构参考 Lenet-5 经典模型[12]的结构, 在特征提取部分, 同样使用两组卷积层和最大池化层最后通过一个全连接层提取图像的抽象特征; 在特征映射部分, 在保留图像目标识别功能的同时, 网络中增加一条分支实现目标方向估计, 即连接一个由两个神经元构成的全连接层, 输出方向分量  $(x, y)$ 。完整网络结构如图 6 所示。

目标识别部分的网络与 Lenet-5 一致, 具体介绍方向估计部分的网络。由于两个输出分量  $(x, y)$  之间存在单个强约束  $\varphi = x^2 + y^2 - 1 = 0$ , 则按照本文第一节所述, 考虑单个输出约束的网络结构如图 7 所示, 其中网络的整个特征提取部分抽象为一个抽象特征层。

在图 7 中, 不仅需要计算最终预测的估计值  $(\tilde{x}, \tilde{y})$  与真值  $(\bar{x}, \bar{y})$  之间的误差, 以  $Diff_\theta$  表示, 则

$$Diff_\theta = (\tilde{x} - \bar{x}) + (\tilde{y} - \bar{y}) \tag{17}$$

还需要考虑输出  $(\tilde{x}, \tilde{y})$  之间的约束关系  $\varphi$ , 计算约束误差, 约束误差函数为  $Error_\varphi$ , 则

$$Error_\varphi = \tilde{\varphi} - 0 = \tilde{x}^2 + \tilde{y}^2 - 1 \tag{18}$$

在利用大量标注有方向的图片样本集对网络进行训练的过程中, 通过误差反向传播对各权值进行修正, 过程示意图如图 8 所示。

权重更新计算推导如下。

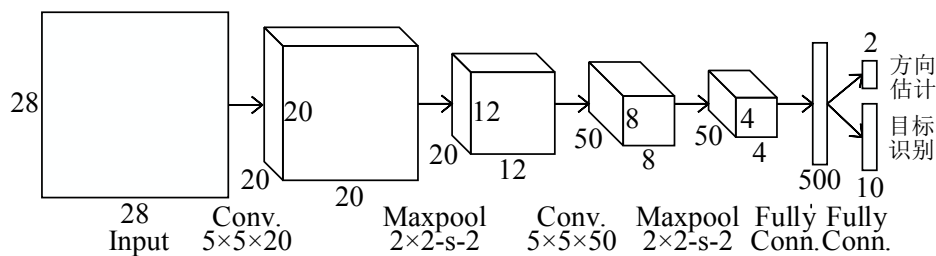


Figure 6. Neural network structure of target recognition and direction estimation based on Lenet-5  
图 6. 基于 Lenet-5 的目标识别和方向估计网络结构图

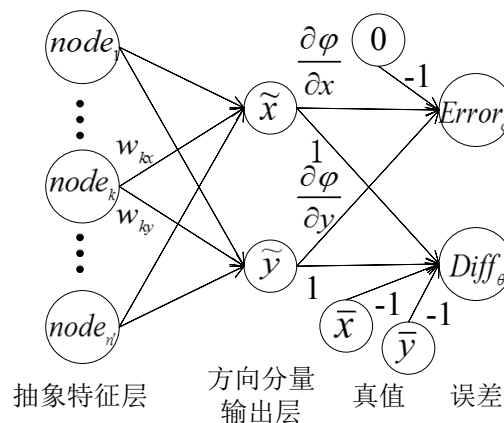
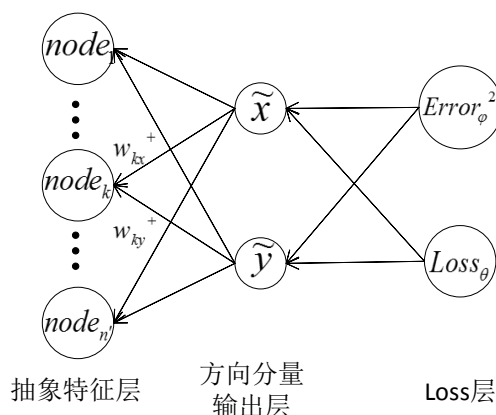


Figure 7. Neural network structure with constraints of directional component outputs  
图 7. 考虑方向分量输出约束的神经网络结构图





**Figure 8.** Back propagation process of the neural network with constraints of directional component outputs  
**图 8.** 考虑方向分量输出约束的神经网络反向传播示意图

Loss 层中，使用均方差定义输出损失函数  $Loss_\theta$ ，则

$$Loss_\theta = \frac{1}{2} \sum [(\tilde{x} - \bar{x})^2 + (\tilde{y} - \bar{y})^2] \quad (19)$$

$Loss_\theta$  在反向传播时，对  $\tilde{x}$  和  $\tilde{y}$  的偏导分别为

$$\begin{cases} \frac{\partial Loss_\theta}{\partial \tilde{x}} = (\tilde{x} - \bar{x}) \\ \frac{\partial Loss_\theta}{\partial \tilde{y}} = (\tilde{y} - \bar{y}) \end{cases} \quad (20)$$

约束误差的平方为

$$Error_\phi^2 = \frac{1}{4}(\phi - 0)^2 = \frac{1}{4}(\tilde{x}^2 + \tilde{y}^2 - 1)^2 \quad (21)$$

其中增加的系数  $\frac{1}{4}$  是为了平衡求偏导引入的系数。

则  $Error_\phi^2$  在反向传播时，对  $\tilde{x}$  和  $\tilde{y}$  的偏导分别为

$$\begin{cases} \frac{\partial Error_\phi^2}{\partial \tilde{x}} = (\tilde{x}^2 + \tilde{y}^2 - 1) \cdot \tilde{x} \\ \frac{\partial Error_\phi^2}{\partial \tilde{y}} = (\tilde{x}^2 + \tilde{y}^2 - 1) \cdot \tilde{y} \end{cases} \quad (22)$$

记总体误差为  $E_{\text{sum}}$ ，则

$$E_{\text{sum}} = \lambda \cdot Error_\phi^2 + Loss_\theta \quad (23)$$

其中  $\lambda$  为约束误差的常数系数，实验中讨论  $\lambda$  的取值。

则  $E_{\text{sum}}$  在反向传播时，对  $\tilde{x}$  和  $\tilde{y}$  的偏导分别为

$$\begin{cases} \frac{\partial E_{\text{sum}}}{\partial \tilde{x}} = \lambda \frac{\partial Error_\phi^2}{\partial \tilde{x}} + \frac{\partial Loss_\theta}{\partial \tilde{x}} \\ \frac{\partial E_{\text{sum}}}{\partial \tilde{y}} = \lambda \frac{\partial Error_\phi^2}{\partial \tilde{y}} + \frac{\partial Loss_\theta}{\partial \tilde{y}} \end{cases} \quad (24)$$

代入公式(20)和(22), 则  $E_{\text{sum}}$  对  $\tilde{x}$  和  $\tilde{y}$  的偏导为

$$\begin{cases} \frac{\partial E_{\text{sum}}}{\partial \tilde{x}} = \lambda \cdot (\tilde{x}^2 + \tilde{y}^2 - 1) \cdot \tilde{x} + (\tilde{x} - \bar{x}) \\ \frac{\partial E_{\text{sum}}}{\partial \tilde{y}} = \lambda \cdot (\tilde{x}^2 + \tilde{y}^2 - 1) \cdot \tilde{y} + (\tilde{y} - \bar{y}) \end{cases} \quad (25)$$

权值修正根据导数的链式法则计算, 例如, 可通过下式更新权值  $w_{kx}$  和  $w_{ky}$ ,

$$\begin{cases} \frac{\partial E_{\text{sum}}}{\partial w_{kx}} = \frac{\partial E_{\text{sum}}}{\partial \tilde{x}} \times \frac{\partial \tilde{x}}{\partial w_{kx}} \\ \frac{\partial E_{\text{sum}}}{\partial w_{ky}} = \frac{\partial E_{\text{sum}}}{\partial \tilde{y}} \times \frac{\partial \tilde{y}}{\partial w_{ky}} \end{cases} \quad (26)$$

$$\begin{cases} w_{kx}^+ = w_{kx} - \eta \cdot \frac{\partial E_{\text{sum}}}{\partial w_{kx}} \\ w_{ky}^+ = w_{ky} - \eta \cdot \frac{\partial E_{\text{sum}}}{\partial w_{ky}} \end{cases} \quad (27)$$

其中  $\eta$  为神经网络的学习率。

其他部分与传统神经网络处理方法一致, 由此对网络中所有权值进行更新迭代, 得到模型最佳参数矢量  $W$ 。

### 3.4. 方向角精度评定

神经网络预测出一组  $(\cos \tilde{\theta}, \sin \tilde{\theta})$ , 可以通过反三角函数公式  $\tilde{\theta} = \arctan \frac{\tilde{x}}{\tilde{y}}$ , 并结合  $\cos \tilde{\theta}$  和  $\sin \tilde{\theta}$  的符号, 唯一确认目标方向预测值  $\tilde{\theta}$ 。若用  $\bar{\theta}$  表示真实值, 那么方向角的估计误差  $\Delta\theta$  为

$$\Delta\theta = \tilde{\theta} - \bar{\theta} \quad (28)$$

而由于  $\theta$  在  $360^\circ$  到  $0^\circ$  时数值有一个跳变, 简单地利用该式计算可能会产生很大偏差。因为单个方向角的  $loss$  可以由公式(19)计算, 将式中的方向分量分别用方向角的正弦和余弦展开, 可以得到

$$loss = \frac{(\sin \tilde{\theta} - \sin \bar{\theta})^2 + (\cos \tilde{\theta} - \cos \bar{\theta})^2}{2} = \frac{\sin^2 \tilde{\theta} + \cos^2 \tilde{\theta} + 1}{2} - \cos(\tilde{\theta} - \bar{\theta}) \quad (29)$$

$$\cos(\Delta\theta) = \frac{\sin^2 \tilde{\theta} + \cos^2 \tilde{\theta} + 1}{2} - loss \quad (30)$$

所以, 估计误差  $\Delta\theta$  可以由  $loss$  和方向分量估计值反算得到, 即

$$|\Delta\theta| = \arccos\left(\frac{\sin^2 \tilde{\theta} + \cos^2 \tilde{\theta} + 1}{2} - loss\right) \quad (31)$$

如果有  $N$  个样本测试, 那么所有测试样本的方向估计均方误差  $E$  则为

$$E = \frac{\sum_{i=1}^N \Delta\theta_i^2}{N} = \frac{\sum_{i=1}^N \arccos^2\left(\frac{\sin^2 \tilde{\theta}_i + \cos^2 \tilde{\theta}_i + 1}{2} - loss_i\right)}{N} \quad (32)$$

其中,  $i=1, 2, \dots, N$ 。

这样, 利用公式(31)和(32)计算得到的估计误差和均方误差不会因为角度值的跳变而出现突变。

公式(32)也反映了方向分量损失与方向估计精度的关系, 当方向分量估计损失越小时, 均方误差  $E$  越小, 方向估计精度越高。

另外方向分量估计值的平方和也会影响估计误差, 由于其存在  $\tilde{\varphi} = \sin^2 \tilde{\theta} + \cos^2 \tilde{\theta} - 1 = 0$  的约束, 还需关注约束误差, 约束误差由公式(18)和(21)计算得到, 约束误差越小, 方向估计精度越高。

## 4. 实验

### 4.1. 实验数据及运行环境

本文对高清航拍影像的飞机目标进行方向估计测试, 原始数据来源于中科院大学高清航拍目标数据集 UCAS-AOD [13]中的包含飞机目标的高清航拍影像, 采用数据集中飞机目标位置标注数据裁剪出目标。整理后, 最终使用 893 张影像内包含的 6447 架飞机作为训练集, 104 张影像内的 1005 架飞机作为测试集, 比例约为 6.4:1。部分实验样本示例图如图 9 所示。

目标方向为从飞机机尾指向机头的方向, 编写方向标定程序对所有高清航拍影像中的飞机目标进行人工方向标注, 首先画出从机尾到机头的直线, 然后程序自动计算所画直线的方向角  $\theta$  的两分量值, 作为训练时的真值, 存储在样本标定文本数据中。

本文全部程序采用 C++语言, 在 Caffe 框架下编写, Visual Studio 2013 编译环境, 硬件采用 Intel i5-4210U 处理器, 1.7 GHz, 8 G 内存, 单线程运行, 无 GPU 加速。

### 4.2. 实验步骤

#### 4.2.1. 训练

使用 UCAS-AOD 数据集中 6447 张飞机影像及相应的目标方向分量值作为学习数据, 在 Caffe 框架下实现图 6 所示的网络结构, 神经网络基础学习率  $\eta = 0.01$ , 分别采用不考虑输出约束的传统方法, 即约束误差常数系数  $\lambda = 0$ , 和本文提出的考虑单个输出约束的方法, 并考虑约束误差常数系数  $\lambda$  分别为 0.5, 1.0, 1.5, 对学习数据进行训练, 迭代 50,000 次, 得到训练模型, 并绘制曲线观察每次迭代方向分量均方差 Loss 值和约束误差值的变化。



Figure 9. Diagram of experimental samples for target direction estimation

图 9. 目标方向估计实验样本示例图

### 4.2.2. 测试

测试：使用 UCAS-AOD 中 1005 张飞机影像及相应的目标方向分量值作为测试数据，使用对应于训练网络的网络结构和训练得到的训练模型对测试数据进行方向估计，并采用方向角绘制程序将最终预测的方向显示展示出来。

### 4.3. 实验结果

训练过程中，当约束误差常数系数  $\lambda$  分别为 0 (即不考虑约束误差的传统神经网络方法)、0.5、1.0 及 1.5 时，目标方向分量的均方差损失和约束误差在前 8000 次迭代中的变化曲线如图 10 和图 11 所示。

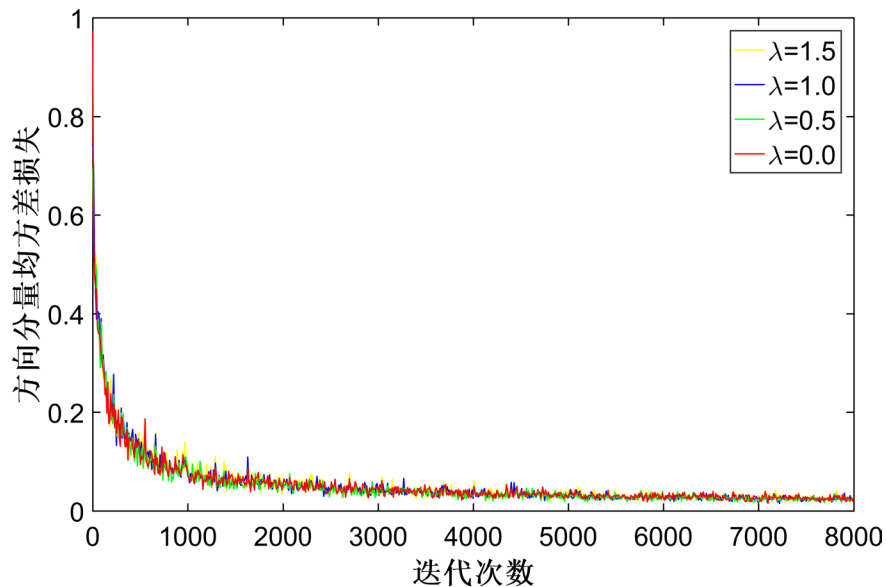


Figure 10. Variation diagram of MSE Loss of target direction components with iterations

图 10. 目标方向分量均方差损失随迭代次数变化示意图

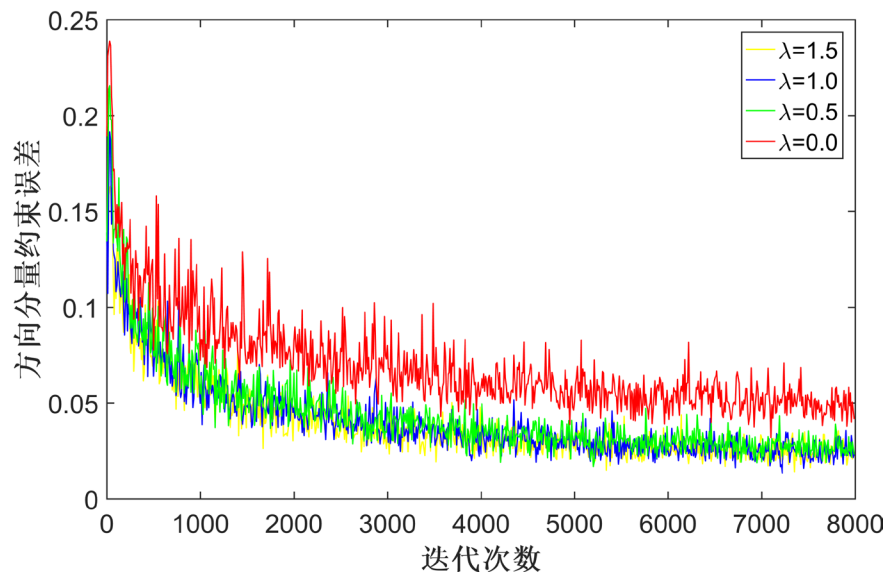


Figure 11. Variation diagram of constraint errors of target direction components with iterations

图 11. 目标方向分量约束误差随迭代次数变化示意图

当约束误差常数系数不为 0(红色曲线)时, 即网络增加了约束误差的考量后, 从图 10 中可以看出, 目标方向分量的均方差损失值随迭代次数增加下降的速度和幅度与不考虑约束误差的传统方法几乎不变, 维持在传统方法的水平; 而从图 11 中可以明显看到, 约束误差随迭代次数下降速度明显更快, 且当约束误差趋于稳定时, 约束性神经网络的方法得到的最终值比没有考虑约束的传统方法低近一倍。

因此, 本文提出的方法, 能提高约束误差下降的速度和幅度。而由公式(31)和(32)可以看出, 减小约束误差, 对减小方向估计误差是有作用的。因此, 运用引入约束误差的约束化神经网络进行目标方向估计能够减小约束误差, 提高方向估计精度。

关于约束误差常数系数  $\lambda$  的选择, 在图 10 和图 11 中  $\lambda$  分别为 0.5, 1.0, 1.5 时, 绿蓝黄三种曲线的走势相似, 在此应用中, 约束误差对整体预测的影响有一定限度, 不过在实验发现, 系数越大时网络越不容易收敛, 所以建议在具体应用中, 不宜使用较大的系数。

图 12 是使用约束误差常数系数  $\lambda$  为 0.1 的约束性神经网络, 迭代 50,000 次的训练模型, 对测试图片进行目标方向估计的部分结果图。

## 5. 结论与展望

本文提出了约束性神经网络的概念, 进一步提出了在具体问题中网络输出分量间存在弱约束和强约束时的一般性处理方法, 提供了考虑约束误差的新思路新方法, 具有普适性和较高的应用价值。在图像目标方向估计的应用中, 本文对描述目标方向的方向角的两个具有强约束关系的方向分量进行回归, 故采用约束性神经网络, 在 Loss 层引入约束误差, 解决目标方向分量估计问题。实验表明, 基于约束性神经网络的目标方向估计方法, 能够在保证原输出损失的下降速度和幅度的前提下, 减小约束误差, 提高方向估计精度。

后续研究工作可以集中于采用结构更成熟的神经网络结构, 实现利用目标的方向信息纠正图像、更加准确地定位任意角度倾斜的细长条状物体等一些具体应用, 使用约束性神经网络解决更多具有输出约束的实际问题, 如预测目标在三维空间内的方向等。

## 致 谢

本文工作在刘进副教授的指导和方高硕士的帮助下完成, 感谢中科院大学高清航拍目标数据集



Figure 12. Diagram of part of target direction estimation results in test images

图 12. 测试图片目标方向估计的部分结果图

UCAS-AOD [13], 感谢国家自然科学基金项目“球面全方位目标测量及跟踪算法研究”(编号 41271454)对本文工作的支持, 特此致谢。

## 基金项目

国家自然科学基金项目, 编号 41271454。

## 参考文献

- [1] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, **1**, 1097-1105.
- [2] Girshick, R., Donahue, J., Darrell, T., et al. (2014) Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 580-587.
- [3] Girshick, R. (2015) Fast r-cnn. arXiv Preprint arXiv:1504.08083.
- [4] Ren, S., He, K., Girshick, R., et al. (2016) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149.
- [5] Redmon, J., Divvala, S., Girshick, R., et al. (2016) You only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 779-788.
- [6] Liu, W., Anguelov, D., Erhan, D., et al. (2016) SSD: Single Shot Multibox Detector. *European Conference on Computer Vision*, Springer, Berlin, 21-37.
- [7] Redmon, J. and Farhadi, A. (2017) YOLO9000: Better, Faster, Stronger. arXiv Preprint.
- [8] Jiang, Y., Zhu, X., Wang, X., et al. (2017) R2CNN: Rotational Region CNN for Orientation Robust Scene Text Detection. arXiv Preprint arXiv:1706.09579.
- [9] Tang, T., Zhou, S., Deng, Z., et al. (2017) Arbitrary-Oriented Vehicle Detection in Aerial Imagery with Single Convolutional Neural Networks. *Remote Sensing*, **9**, 1170. <https://doi.org/10.3390/rs9111170>
- [10] Liu, L., Pan, Z. and Lei, B. (2017) Learning a Rotation Invariant Detector with Rotatable Bounding Box. arXiv Preprint arXiv:1711.09405.
- [11] Nasrabadi, N.M. (2007) Pattern Recognition and Machine Learning. *Journal of Electronic Imaging*, **16**, 049901. <https://doi.org/10.1117/1.2819119>
- [12] LeCun, Y. (2015) LeNet-5, Convolutional Neural Networks. <http://yann.lecun.com/exdb/lenet>
- [13] Zhu, H., Chen, X., Dai, W., et al. (2015) Orientation Robust Object Detection in Aerial Images Using Deep Convolutional Neural Network. 2015 *IEEE International Conference on Image Processing (ICIP)*, Quebec City, 27-30 September 2015, 3735-3739.

### 知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2329-549X, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [gst@hanspub.org](mailto:gst@hanspub.org)