

基于深度神经网络剪枝的两阶段遥感图像目标检测

王生生, 王 萌, 王光耀

(吉林大学 计算机科学与技术学院, 吉林 长春 130012)

摘 要: 在高分辨率遥感图像目标检测中,受云雾、光照、复杂背景、噪声等因素影响,现有目标检测方法虚警率高、速度慢、精确度低.为此提出基于深度神经网络剪枝的两阶段目标检测(object detection based on deep pruning, ODDP)方法.首先,给出深度神经网络剪枝方法,基于深度神经网络剪枝分别提出自主学习区域提取网络算法与优化训练分类网络算法;然后,将上述两算法用于卷积神经网络,得到两阶段目标检测模型.实验结果表明,在NWPU VHR-10高分辨率遥感数据集上,相比现有目标检测方法,ODDP的检测速度和精度均有一定提升.

关 键 词: 计算机视觉;目标检测;高分辨率遥感图像;深度学习;卷积神经网络

中图分类号: TP 391

文献标志码: A

文章编号: 1005-3026(2019)02-0174-06

Deep Neural Network Pruning Based Two-Stage Remote Sensing Image Object Detection

WANG Sheng-sheng, WANG Meng, WANG Guang-yao

(College of Computer Science and Technology, Jilin University, Changchun 130012, China. Corresponding author: WANG Meng, E-mail: 1030997649@qq.com)

Abstract: In the object detection of high-resolution remote-sensing images, affected by cloud, light, complex background, noise and other factors, the existing object detection method has high false alarm, low speed and low precision. So we propose a two-stage object detection method based on deep pruning. First, we propose deep pruning, and then based on the deep pruning we propose an algorithm that learns region proposal network automatically and an algorithm that we train classification network with optimizing training method. We then apply the two algorithms to convolutional neural network and get a two-stage object detection model. The experiment result shows that our method has a certain improvement on precision and speed compared with the state-of-the-art method.

Key words: computer vision; object detection; high-resolution remote sensing image; deep learning; convolutional neural network

光学遥感传感器空间分辨率在过去十年中有了很大的提高,大量高分辨率图像被用于军事和民用领域.但由于卫星拍摄时受云层、雾霾、光照不均匀等天气影响,会引入大量噪声,而且往往遥感图像的背景较为复杂,基于SIFT, HOG等传统目标检测方法处理这类图像时效果并不好.近来使用深度卷积神经网络的两阶段目标检测方法被应用于高分辨率遥感图像目标检测.两阶段即区域提取(region proposal)阶段^[1]和分类(classification)

阶段.其中,区域提取阶段要求能快速提取出候选区域,分类阶段要求能准确分类候选区域属于哪一个类别.

使用深度卷积神经网络的两阶段目标检测方法在高分辨率遥感图像的舰船、飞机、车辆等单目标、多目标检测中效果显著. Long等^[2]使用selective search(SS)算法生成候选区域,通过卷积神经网络CNN(convolutional neural network)和SVM分类器对光学遥感图像中的多目标进行

检测. Jiang 等^[3]使用基于图的超像素分割方法提取图像碎片,将遥感图像碎片输入 CNN 进行分类,检测碎片是否属于车辆. Ševo 等^[4]直接将高分辨率卫星图像切割为相同大小的图像碎片,将图像碎片输入训练好的 CNN 进行分类. Wu 等^[5]利用遥感图像中目标的边缘信息,使用 Edge Boxes 提取候选区域,并将候选区域输入 CNN 进行分类. 但使用 SS、超像素分割方法提取候选区域速度慢,而且直接切割会损失图像信息;Edge Boxes 对复杂背景图像的候选区域提取效果不理想,而且在分类阶段要使用一次训练卷积神经网络. 总之,现有深度卷积神经网络两阶段目标检测方法在面对复杂背景的高分辨率遥感图像时存在虚警率高、速度慢等问题.

为解决上述问题,本文提出基于深度神经网络剪枝的两阶段遥感图像目标检测方法(object detection based on deep pruning, ODDP). 首先提出深度神经网络剪枝,然后基于该方法提出自主学习区域提取网络算法(learning region proposal network, LRPN). 最后提出基于深度神经网络剪枝的优化训练分类网络(optimizing classification network, OCN)算法,学习多种目标地物不同特征,以提高分类网络精度. 将区域提取和分类两个深度卷积神经网络相结合,建立两阶段目标检测模型,即可对高分辨率遥感图像进行检测.

1 基于深度神经网络剪枝的两阶段目标检测方法

为解决高分辨率遥感图像目标检测中虚警率低、速度慢、精度低的问题,本文在提出深度神经网络剪枝的基础上,又提出了基于深度神经网络剪枝的两阶段遥感图像目标检测方法,包括自主学习区域提取网络算法 LRPN 和优化分类网络算法 OCN. LRPN 通过两次剪枝,获得区域提取网络. OCN 利用 LRPN 输出的区域提取网络,获得分类网络训练集,训练时利用剪枝方法获得训练好的分类网络. 将区域提取网络和分类网络结合,即获得两阶段目标检测模型.

1.1 深度神经网络剪枝

深度神经网络通常过参数化(over-parameterized),存在显著的冗余性^[6],导致网络在拟合小模型(即参数少的函数)时,可能过拟合,计算速度慢. 受神经网络的丢弃(dropout)方法启发,本研究提出深度神经网络剪枝(deep pruning, DP)算法,可减小网络容量,降低过拟合

概率,加速网络计算. 具体来说,首先计算出核的度量值,然后剪去低于度量值的核(filter),从而改变网络的结构,使网络更稀疏.

深度神经网络剪枝分为六步:

①训练深度神经网络.

②对第 i 层中的第 j 个核 $F_{i,j}$, 计算其所有权重的绝对值之和 S_j :

$$S_j = \sum_1^{n_i} \sum_1^{\text{size}} |w| . \quad (1)$$

式中: n_i 表示第 i 层输入通道数; size 表示第 j 个核的大小; w 表示权重.

③对以上计算的所有 S_j 进行升序排序.

④根据排序结果,剪去前 m_i 个核,并剪去这 m_i 个核对应的特征图.

⑤在第 i 层剪完后,对新的核矩阵进行训练.

⑥分析剪枝前后的网络平均精确度差值 ΔP , 若 $\Delta P < 0$, 则 $m_i = m_i - 1$, 返回步骤④重新剪枝, 否则返回步骤②对第 $i + 1$ 层进行相同操作.

对于网络的全连接部分,将其看作 1×1 大小的核,从而可使用上述方法剪枝.

1.2 自主学习区域提取网络

现有技术 in 高分辨率遥感图像的区域提取阶段使用 SS、Edge Boxes 等基于手工设计特征的方法; 在面对背景干净、目标特征易表达的图像时, 这些方法表现很好, 但在面对复杂背景图像时, 区域提取效果差. 针对此问题, 本文提出自主学习区域提取网络(learning region proposal network, LRPN)算法.

LRPN 算法使初始网络从数据集中自主学习构建一个具有新结构的卷积神经网络. 自主学习的卷积网络有以下优点:

①网络结构符合数据集分布;

②既具有初始网络的整体结构又更稀疏;

③避免人工设计网络结构的盲目性, 可靠性高. 因而 LRPN 算法计算速度快, 且不影响目标检测的精度. 下文讲述 LRPN 算法及其伪代码.

首先对初始网络进行预训练, 为后续训练提供初级特征, 若后续训练集小, 使用预训练可收敛更快. 预训练采用遥感数据集 AID(aerial image dataset)^[7], 而不采用 ImageNet 数据集, 因此预训练提取的遥感特征更加适合于遥感任务. 为从初始网络自主学习构建新网络结构, 使用深度神经网络剪枝作为一种改变初始网络结构的方式. 在第一次剪枝时, 减去对二分类不重要的核, 为后续步骤提供卷积层权重, 并设置较小稀疏度, 为定位任务留下空间.

第一次剪枝后对网络进行重构,重构时选取网络的卷积层,在其后加入两个分支:一支为可区分背景与目标的二分类的 cls_fc_class ,另一支为提供坐标偏移的 cls_fc_bbr . 训练时会将 cls_fc_class 输入到 softmax 分类器中,分类器给出图像为背景或目标的概率; cls_fc_bbr 在训练时用 L1 损失函数进行回归训练. 重构可使网络具有同时处理二分类和定位的能力.

然后进行第二次剪枝,去掉对定位、二分类都冗余的核. 第二次剪枝和第一次剪枝方式相同,两次剪枝稀疏度不同.

第二次剪枝再一次改变网络结构,使网络结构符合数据集的分布. 首先对重构后的网络进行训练,训练任务是二分类(背景与目标)以及定位,训练目的是判断哪些核不仅对二分类有效,而且对定位有效. 因有两个任务,相应地,目标函数需采用多任务损失函数. 同时优化目标、背景二分类和定位能力的多任务损失函数 L_{ODDP} 为

$$L_{\text{ODDP}}(\text{loc}, p^{\text{bic}}) = L_{\text{bic}}(p^{\text{bic}}) + \alpha L_{\text{bbox}}(\text{loc}). \tag{2}$$

式中: loc 是候选区域坐标; p^{bic} 为分类器算出的类别置信度; L_{bic} 是目标、背景二分类 softmax 损失; loc_i 是目标的坐标标签; L_{bbox} 是 smooth L1 损失.

$$L_{\text{bbox}}(\text{loc}) = f_{\text{Li}}(\text{loc} - \text{loc}_i); \left. \begin{aligned} f_{\text{Li}}(x) = \begin{cases} 0.5x^2, & |x| < 1; \\ |x| - 0.5, & \text{其他.} \end{cases} \end{aligned} \right\} \tag{3}$$

在反向传播中背景的包围盒无意义,而且会使模型过早收敛. 因此,设背景样例 $\alpha = 0$, 目标样例 $\alpha = 0.5$.

最后使用区域提取训练数据集与其类别标签和坐标标签,对网络进行二分类和定位的训练.

以上训练完成以后就可使用最终网络的卷积层构建区域提取网络 RPN (region proposal network),所以区域提取网络其实是全卷积神经网络.

LVPN 算法伪代码如下:

Input: 初始网络 N ; 预训练数据集 D_{pretrain} 及其标签 L_{pretrain} ; 区域提取训练数据集 D_{rpn} 及其标签 $L_{\text{class}}, L_{\text{bbx}}$

Output: 区域提取网络 N_{rpn}

步骤 1 预训练,其中 TRAIN_NET 为训练网络的函数:

- ①初始化 N 的权重 $W, W_{(0)} \sim N(0, 0.01^2)$;
- ② $W_{\text{pretrain}} = \text{TRAIN_NET}(W_{(0)}, D_{\text{pretrain}}, L_{\text{pretrain}})$.

步骤 2 第一次剪枝:

对 W_{pretrain} 应用 DP 算法得到新的网络结构 $W_{\text{pruned-one}}$,其中数据集使用区域提取训练数据集的 D_{rpn} 及其标签 L_{class} .

步骤 3 重构:

使用 $W_{\text{pruned-one}}$ 中的卷积层及其权重 W_{conv} ,组合两分支 $\text{conv_fc_class}(\text{softmax})$ 与 conv_fc_bbr (L1 回归),得到 N_{rpn} .

步骤 4 第二次剪枝:

- ①将 N_{rpn} 中新加入的核 W_{new} 初始化, $W_{\text{new}} \sim N(0, 0.01^2)$;
- ② $W_{(0)} = (W_{\text{conv}}, W_{\text{new}})$;
- ③对 $W_{(0)}$ 使用 DP 算法得到新的网络结构 $W_{\text{pruned-two}}$,其中数据集使用区域提取训练数据集的 D_{rpn} 及其标签 $L_{\text{class}}, L_{\text{bbx}}$;
- ④选取 N_{rpn} 卷积层,获得 N_{rpn} .

1.3 优化训练分类网络

复杂网络的好处是非常具有表现力,可以捕捉到特征与输出之间高度非线性关系. 这种大型网络的缺点是倾向于捕捉训练数据集中的噪声,这种噪声并不能推广到新的数据集中,导致过拟合和高方差,泛化能力弱;而简单地减少模型容量会导致另一个极端,网络会错过特征和输出之间的相关关系,从而导致欠拟合和高偏差. 现有方法在训练分类网络时,简单地随机选取深度神经网络进行训练,数据集与网络容量并不能很好匹配,从而造成过拟合或欠拟合. 本文提出了一种基于深度神经网络剪枝的优化训练算法,在剪枝与核恢复中调节网络,使网络更好地学习数据集的分布. 下文详细讲述优化训练分类网络算法 (optimizing classification network, OCN) 及 OCN 的伪代码.

在 OCN 算法中,使用区域提取网络生成的分类网络训练集对分类网络进行训练,可使整个目标检测模型保持一致性. 生成分类网络训练集的步骤是:给训练集输入区域提取网络;产生候选区域;按照正:负 = 1:3 的比例选取候选区域集. 本研究定义 IoU (intersection over union) ≥ 0.7 的图片块为正样例, $\text{IoU} < 0.3$ 的为负样例. IoU 是目标检测中的一个评价指标,是模型检测出的候选区域与坐标标签区域的交叠率:

$$\text{IoU} = \frac{G \cap D}{G \cup D}. \tag{4}$$

式中: D 是模型检测出的某个候选区域的窗口; G 是待检测图像的标签区域.

为和区域提取网络取得一致性,采用 LVPN 输出的 W_{pretrain} 作为 OCN 的输入网络. 首先对

W_{pretrain} 训练, 然后对网络进行 DP 剪枝. 与区域提取网络相比, 分类网络需学习多类别中每个类别具有的专属特征, 并需区分每个类别之间的细微差异, 因此需更大的模型容量, 表现在网络上就是使用更多的核. 因此在对网络进行剪枝再训练后, 恢复了剪掉的核. 最后在训练网络时, 因剪枝后的网络经过再训练已经到达一个较好的局部最小值, 所以这一步训练时学习率设为上一步的 $1/10$. 通过恢复剪掉的核和训练, 模型容量增加, 再训练后网络可能到达一个更好的局部最小值.

OCN 算法伪代码如下:

Input: 分类网络训练集 D_{cn} 及其标签 L_{class}

Output: N_{cn}

步骤 1 剪枝:

①用 LRPN 中的 W_{pretrain} 初始化 $W_{(0)}$;

②对 $W_{(0)}$ 使用 DP 算法得到新的网络结构 W_{pruned} , 其中数据集使用分类网络训练集 $D_{\text{cn}}, L_{\text{class}}$.

步骤 2 恢复核:

①恢复剪掉的核 W_{recover} , 并将 W_{recover} 初始化;

② $W_{(0)} = (W_{\text{pruned}}, W_{\text{recover}})$;

③ $W_{\text{max-iteration}} = \text{TRAIN_NET}(W_{(0)}, D_{\text{cn}}, L_{\text{class}})$;

④ $W_{\text{max-iteration}}$ 即为 N_{cn} 的权重.

2 实验与分析

为评估 ODDP 高分辨率遥感图像目标检测方法, 选用公开数据集 NWPU VHR-10, 这是一个 10 类别高分辨率遥感图像目标检测数据集. 10 类别包括飞机、船舶等(编号 1~10). 在本实验中, 只使用数据集的正样例, 包括 650 张高分辨率光学遥感图像; 其中部分图片来自于谷歌地图, 其空间分辨率在 0.5~2 m 之间, 另一部分来自 Vaihinge 数据集, 其空间分辨率为 0.08 m. 实验中, 选取数据集的 50% 为训练集, 20% 为验证集, 30% 为测试集. 采用精确度-召回率曲线、平均运行时间、平均精确度(average precision, AP) 和 mAP(mean average precision) 来定量评估 ODDP.

在实验中, 初始网络选取 AlexNet; 预训练集选择 AID; 最大迭代次数设为 80 000. 定义 TP 为 $\text{IoU} \geq 0.7$ 且被判别为目标; FN 为未检测到的目标; FP 为 $\text{IoU} < 0.3$ 且被判别为目标. 实验硬件环境是 2.8GHz, 6 核 CPU, 32GB 内存, GPU 型号为 GTX Titan X.

图 1 是 ODDP 检测出的部分图片的结果. ODDP 算法将检测出的目标圈出, 并标注目标区域所属类别及该区域属于此类别的概率. 图 2b 显示了一个误检的假正区域, 图 2c 显示了一个漏检的假负区域. 可以看出, 尽管被检测物体在大小、形状、纹理上有很大不同, 背景也较为复杂, ODDP 仍然检测出大多数目标. 为了定量评估 ODDP, 本文和 4 种目前最优的方法对比, 包括 COPD^[8]、迁移 CNN^[9]、无微调 RICNN^[10]、微调 RICNN^[10]. 4 种方法都采用 SS 提取候选区域. 为公平对比, ODDP 与 4 种方法的对比实验均采用相同训练、测试集. 除 COPD 外, 其余 3 种方法均采用 AlexNet 网络结构.

表 1 和图 2 是定量对比结果, 由实验结果可看出, ODDP 在 10 个类别的平均精度上, 比其他 4 种方法的表现更好. 和微调后 RICNN 对比, ODDP 检测出的 10 类别 AP 分别提高了 0.054 4, 0.129 4, 0.012 0, 0.049 3, 0.155 3, 0.329 5, 0.032 7, 0.057 2, 0.195 3, 0.088 8, mAP 提高了 0.112 6, 表明 ODDP 的检测能力较强. 通过图 2 可以看出, 随着召回率的增加, ODDP 精确度曲线下降得最慢, 说明 ODDP 候选区域中正样例占比大, 因此区域提取网络比 SS 更好.

用平均运行时间评估各个目标检测方法的速度. ODDP 平均运行时间为 0.38 s; COPD, 迁移 CNN, 无微调 RICNN 和微调 RICNN 的平均运行时间分别为 1.07, 5.24, 8.77 和 8.77 s. 由此看出, ODDP 运行时间最短.

表 1 检测方法的平均精度
Table 1 Average precision of the detection methods %

目标	COPD	迁移 CNN	无微调 RICNN	微调 RICNN	ODDP
飞机	60.44	64.48	84.33	85.39	90.83
船	67.85	55.71	74.53	75.38	88.32
储油罐	60.31	81.38	66.87	84.95	86.15
棒球场	81.20	79.78	85.28	85.69	90.62
网球场	52.23	56.46	63.82	65.70	81.23
篮球场	36.71	46.73	57.18	57.96	90.91
田径场	82.71	78.26	82.77	86.32	89.59
海港	82.72	78.26	82.79	84.12	89.84
桥	17.13	45.97	58.25	61.28	80.81
车辆	44.09	43.01	66.56	69.88	78.76
mAP	58.54	63.00	72.24	75.45	86.71

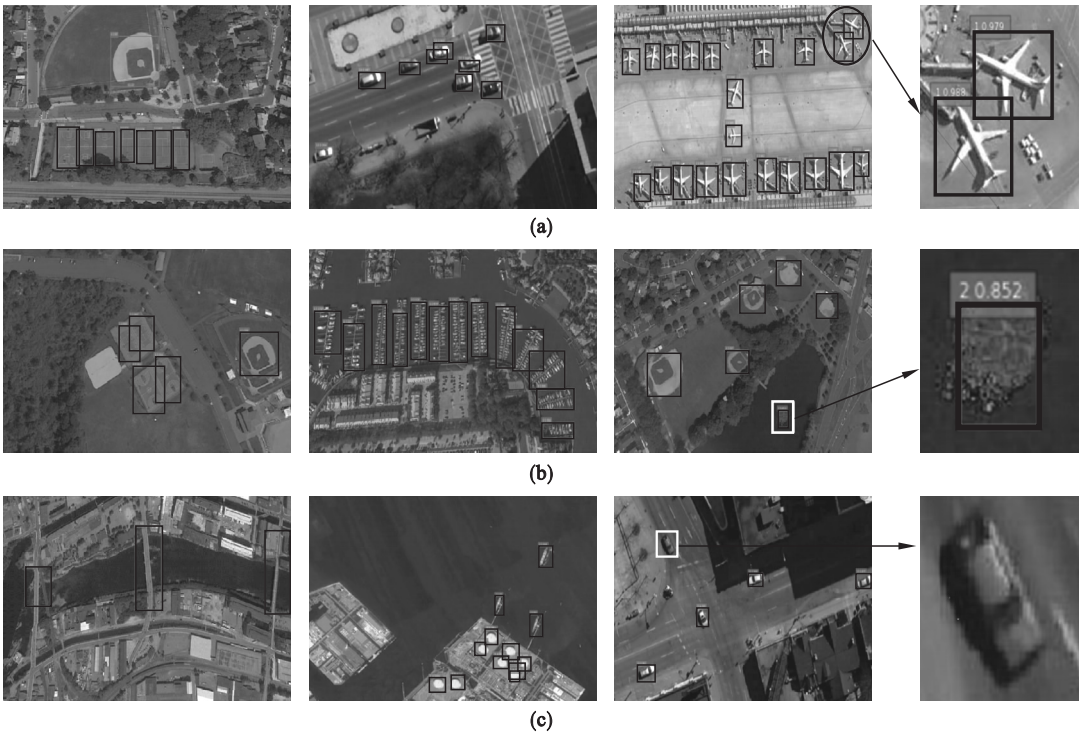


图 1 ODDP 检测结果
Fig. 1 ODDP detection results

(a)—检测出的目标; (b)—误检为假正的区域; (c)—漏检的假负区域.

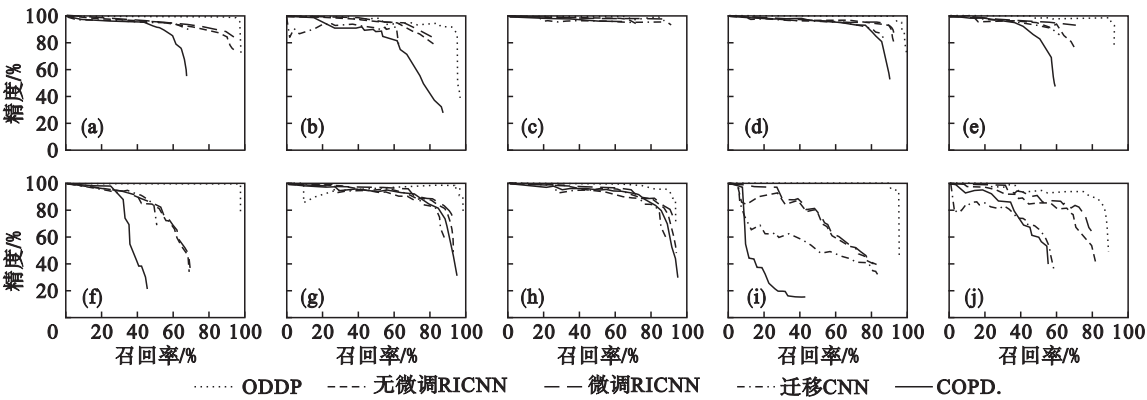


图 2 十类别 PRC 曲线图
Fig. 2 Ten-class PRC curve graphs

(a)—飞机; (b)—船; (c)—储油罐; (d)—棒球场; (e)—网球场; (f)—篮球场; (g)—田径场; (h)—海港;
(i)—桥; (j)—车辆.

3 结 语

本文提出了基于深度神经网络剪枝的两阶段高分辨率遥感图像目标检测方法 ODDP. 受神经网络的丢弃(dropout)方法启发,首先提出深度神经网络剪枝方法 DP,旨在减小网络容量,降低过拟合概率,加速网络计算;然后提出基于 DP 的目标检测训练算法,包括自主学习区域提取算法 LRPN 和优化训练分类网络算法 OCN. 将初始网络输入 LRPN 即可得到区域提取网络. 根据

OCN,将候选区域用于优化训练卷积神经网络,从而得到分类阶段的分类网络. 由区域提取网络和分类网络结合成两阶段目标检测模型,因此模型在两阶段均采用深度 CNN 特征. 实验结果表明,在遥感数据集上 ODDP 比现有的目标检测方法速度更快,精度更好.

参考文献:

[1] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway:IEEE,

-

with genetic algorithms [J]. *Computers & Industrial Engineering*, 2017, 107: 39 – 56.

[5] Boehm B. Anchoring the software process [J]. *IEEE Software*, 1996, 13(4): 73 – 82.

[6] 周方明, 张明媛, 袁永博. 基于 PCA – GA – BP 的工程项目工期风险预测研究 [J]. *工程管理学报*, 2011, 25 (5): 534 – 538.

(Zhou Fang-ming, Zhang Ming-yuan, Yuan Yong-bo. Risk of project time based on PCA-GA-BP [J]. *Journal of Engineering Management*, 2011, 25(5): 534 – 538.)

[7] Dao S D, Abhary K, Marian R. A bibliometric analysis of genetic algorithms throughout the history [J]. *Computers & Industrial Engineering*, 2017, 110: 395 – 403.