

一种改进的卷积神经网络的数显仪表识别方法

李记花, 李鹤喜, 李威龙

五邑大学智能制造学部, 广东 江门

Email: 1073434853@qq.com

收稿日期: 2021年1月7日; 录用日期: 2021年2月2日; 发布日期: 2021年2月8日

摘要

为了提高数显仪表的识别率, 设计了一种传统图像处理方法和深度学习技术相结合的算法, 即一种基于改进的卷积神经网络的数显仪表识别算法。首先通过传统图像处理技术对图像进行图像预处理、字符分割等操作, 然后由基于注意机制的卷积神经网络算法对字符进行识别。实验结果表明, 该方法不仅有效提高了字符的准确率, 字符识别率高达98.5%, 还提高了网络的收敛速度。该方法基本可以满足各种数显仪表的识别, 能够满足实际应用的需求。

关键词

数显仪表识别, 注意机制, 卷积神经网络

An Improved Digital Display Instrument Recognition Method Based on Convolutional Neural Network

Jihua Li, Hexi Li, Weilong Li

Faculty of Intelligent Manufacturing, Wuyi University, Jiangmen Guangdong

Email: 1073434853@qq.com

Received: Jan. 7th, 2021; accepted: Feb. 2nd, 2021; published: Feb. 8th, 2021

Abstract

In order to improve the recognition rate of digital display instruments, an algorithm combining traditional image processing method and deep learning technology is designed, that is, an algorithm for digital display instrument recognition based on an improved convolutional neural network. Firstly, the traditional image processing technology is used to perform image preprocessing,

character segmentation and other operations, and then the characters are recognized by the convolutional neural network algorithm based on the attention mechanism. The experimental results show that this method not only improves the character accuracy effectively, the character recognition rate is up to 98.5%, but also improves the convergence speed of the network. This method can basically meet the recognition of various digital display instruments and meet the requirements of practical application.

Keywords

Digital Display Instrument Identification, Attention Mechanism, Convolutional Neural Network

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

1. 引言

数显仪表，就是一种显示数字的仪器，便于人们了解相关信息。目前，数显仪表被广泛的应用于航天、农业、工业等各个行业中，但出于工作条件和成本控制等原因，仍有很多的仪表无法直接获得读数，大多由人工读取。但是人工无法长时间且实时记录，还有些地方工人不方便记录，这些都会对后期工作产生不利影响。随着科学技术的发展，计算机视觉技术可以通过相机拍摄的图片获取有效信息[1]。数显仪表的识别便可以先通过相机对仪表进行拍摄获取图片，接着对图片进行图像处理操作，然后经过图像识别算法得到数值信息。

基于计算机视觉的仪表识别主要包括三个部分：图像处理、图像分割和图像识别。但图像识别是最关键的部分。传统的图像识别是通过图像的颜色、纹理和形状来提取特征进行识别，是一个非常困难且复杂的工程。目前，随着深度学习技术的发展，基于卷积神经网络的图像识别技术得到了广泛的应用，数显仪表识别的普适性和泛化能力也有了较大提高。视觉注意机制能快速找到图像的显著区域，忽略背景信息，可以提高网络的识别速度与精度。因此，本文提出了一种基于传统图像处理与基于注意机制的卷积神经网络数显仪表识别方法。

2. 相关研究

数字仪表的识别最初是使用模糊识别方法，利用最大隶属原构造数字识别器，是由张海波等人[2]提出。王臣[3]等提出了一种基于模板匹配的最大稳定极值区域定位分割算法 TM-MSER，首先利用人工制作模板去匹配定位出原图像中的数显区域，然后再通过二值化、去噪等图像处理方法对数显区域进一步处理，最后利用投影法对每个字符进行分割，并对分割完成的字符进行识别。这种方法比较复杂，匹配过程比较耗时，且正确率不高。申小阳[4]等提出一种穿线识别的方法，首先基于七段数显表字符的每一段特征建立码表，然后根据每个字符的码表进行查找匹配识别。该方法虽识别简单、速度快但要求数字不倾斜，实际利用率不高。陆靖滨[5]等针对仪表中数显仪表图像的点亮区域多为红色，不亮部分偏灰白的特征设计了一种基于 RGB 彩色空间滤波的数字式仪表读数识别算法，首先使用改进的最大类间方差法完成了对字符区域自适应提取，然后采用基于结构特征的穿线法实现仪表的自动读数识别，该算法虽然能够保留数字区域的亮度特征，但是并不适用于其他类型的数显仪表，普遍适应性较差。蔡梦倩[6]等提出了一种基于全卷积神经网络的数字仪表识别算法，该算法是通过全卷积神经网络实现了图到图的像素级

预测，忽略了图像的预处理过程，此外通过融合加权融合全卷积网络中的多尺度特征和多层级特征，实现数字的识别分类。但是该方法需要海量数据进行训练，前期工作繁琐。

3. 方法概述

本文提出一种基于传统图像处理与基于注意机制的卷积神经网络数显仪表识别方法。方法总体思想如下：首先通过灰度图像、高斯滤波、二值化和形态学处理对图像进行预处理，然后通过轮廓检测算法[7]可找到数显表和每个数字的轮廓，接着通过轮廓质心可对数字进行排序且根据轮廓面积大小比较可找到小数点的位置，并对数字进行分割。最后采用基于注意机制的卷积神经网络算法对数字进行识别得出读数。整个算法流程图如图1所示。

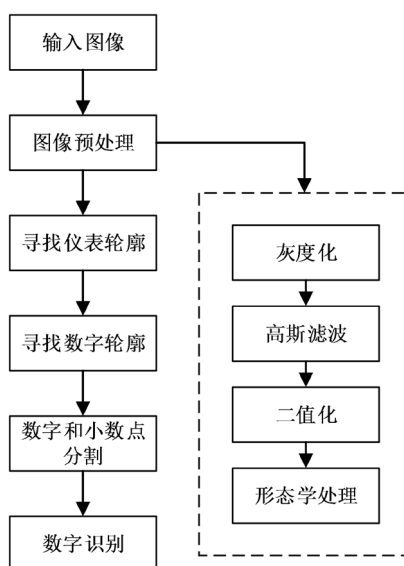


Figure 1. Flow chart of digital display instrument recognition algorithm

图1. 数显仪表识别算法流程图

4. 数显仪表图像预处理

4.1. 灰度转化

图像需要进行灰度转化操作主要是摄像头所采集到的图像是彩色的，这些彩色图像有着大量且复杂的颜色数据，不仅会对后续的图像分割和图像识别造成影响，还占据了较大的存储空间。图像灰度化是将彩色图片中的有效信息提取、只保留亮度信息的过程，这个过程不仅没有丢失卡尺图像的数字信息，反而简化了图像的处理过程，为后续的数字识别提供了方便[8]。图像灰度转化算法有最大值法、平均值法和加权平均值法这三种，前面两种方法考虑因素较少效果不如第三种，因此本文采用加权平均值法来对数显仪表图像进行灰度转化。加权平均值法考虑到人眼对不同颜色的敏感度不一样，从而给 R 、 G 、 B 三个颜色分量匹配不一样的权值。计算表达式为：

$$f(x, y) = W_R R(x, y) + W_G G(x, y) + W_B B(x, y) \quad (1)$$

其中， (x, y) 为每个像素点的坐标； $f(x, y)$ 为每个像素点的灰度结果； R 、 G 、 B 分别为红绿蓝三色通道数据； W_R 、 W_G 、 W_B 为不同颜色分量的权值，一般分别为 0.229、0.587、0.114。

4.2. 高斯滤波

摄像机所拍的数显图像其实有各种各样的噪声，且大多数噪声均为高斯噪声，对图像进行滤波可以去除噪声以便更好的进行后续的图像处理操作和数显仪表轮廓寻找。高斯滤波是一种线性平滑滤波，是图像处理中常用的去噪算法之一，与中值滤波、均值滤波相比平滑效果更柔和，图像边缘也保留较好。高斯滤波首先是用一个模板(卷积、掩模)扫描图像中的每一个像素，然后用模板确定某一点领域内像素的加权平均值，最后用这个平均值去代替模板中心像素点的值，以此达到去噪目的。其计算公式为：

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (2)$$

其中， (x,y) 为像素点的坐标； $G(x,y)$ 为像素点 (x,y) 处的灰度值； σ 为标准差。

4.3. 二值化

图像经过滤波处理后，表盘和数字会呈现不同的灰度值。二值化就是将图像像素点的灰度值设为 255 或 0，图像的颜色便只有黑和白，即数字和表盘背景完全不同色，这样能够有效的将数字从表盘背景中分割出来。二值化表达式为：

$$f(x) = \begin{cases} 255 & \text{if } x \geq T \\ 0 & \text{if } x < T \end{cases} \quad (3)$$

其中， x 为各像素点的灰度值； T 为合适的阈值。

4.4. 形态学处理

形态学处理是图像处理技术中应用最为广泛的一种，主要是为了简化图像数据，可以除掉不相干的结构。在经过二值化处理后，提取的图像中除了数字可能还会存在其它的目标或者数字存在着粘连或断裂的问题，为了区分每个数字和小数点，均需要对其特征进行计算，便可运用形态学技术进行处理。形态学运算指先进行腐蚀运算，平滑图像边界，去除图像中细小连接点，再进行膨胀运算，对二值化图像数字中存在的细小断裂点进行连接。

腐蚀是消除边界点，从而使图像的边界往内部收缩，可以用来消除比较小且没有意义的部分，是利用结构元素来进行平移和填充。定义式为：

$$A \ominus B = \{(x,y) | B_{xy} \cap A \neq \varnothing\} \quad (4)$$

膨胀是一个合并的过程，将与目标有接触的点与目标进行合并成为一个整体，使图像边界向外扩增，可以对目标区域的孔洞进行填充。定义式为：

$$A \oplus B = \{(x,y) | B_{xy} \subseteq A\} \quad (5)$$

其中， A 为原始图像， B 为结构元素。

5. 寻找仪表和数字轮廓

图像的轮廓一般是灰度值变化很大的地方，通过轮廓提取算法可以将仪表框和每个数字提取出来，便于后期的数字分割。这里采用 findContours 算法进行轮廓提取。该算法首先选取某一像素点作为中心，然后以此为中心进行相似性评估，并对周围的像素点进行判断，将符合条件的像素点划分为同一个区域，同时对符合情况像素点周围的点进行类似的判断，最后将全部的像素区域连接。其计算流程如下：假设输入的图像为 $F = \{f_{ij}\}$ ，

如果 $f_{ij} = 1$ 且 $f_{i,j-1} = 0$, 则 (i, j) 是外边界开始点: 以 (i, j) 为中心, (i_2, j_2) 为起始点, 按顺时针方向查找 (i, j) 的邻域是否存在非 0 像素点。若找到非 0 像素点, 则令 (i_1, j_1) 是顺时针方向的第一个非 0 像素点; 否则令 $f_{ij} = -NBD$, 转到(11)。

$$NBD+ = 1, (i_2, j_2) \leftarrow (i, j-1) \quad (6)$$

此时,

$$(i_2, j_2) \leftarrow (i_1, j_1), (i_3, j_3) \leftarrow (i, j) \quad (7)$$

接着以 (i_3, j_3) 为中心, 按逆时针方向, (i_2, j_2) 的下一个点为起始点查找 (i_3, j_3) 的邻域是否存在非 0 像素点, 令 (i_4, j_4) 是逆时针方向的第一个非 0 像素点。如果 $(i_3, j_3 + 1)$ 是已经检查过的像素点且是 0 像素点, 则 $f_{i_3, j_3} \leftarrow -NBD$ 。如果 $(i_3, j_3 + 1)$ 不是已经检查过的像素点且是 0 像素点, 则 $f_{i_3, j_3} \leftarrow NBD$ 。其他情况, f_{i_3, j_3} 不变。如果 $(i_4, j_4) = (i, j)$ 且 $(i_3, j_3) \leftarrow (i_1, j_1)$, 则转到(11); 否则令

$$(i_2, j_2) \leftarrow (i_3, j_3), (i_3, j_3) \leftarrow (i_4, j_4) \quad (8)$$

如果 $f_{ij} \geq 1$ 且 $f_{i,j+1} = 0$, 则 (i, j) 是孔边界开始点:

$$NBD+ = 1, (i_2, j_2) \leftarrow (i, j+1) \quad (9)$$

若 $f_{ij} > 1$, 则

$$LNBD \leftarrow f_{ij} \quad (10)$$

如果 $f_{ij} \neq 1$, 则

$$LNBD \leftarrow |f_{ij}| \quad (11)$$

从点 $(i, j+1)$ 继续光栅扫描, 扫描到图像的右下角的顶点时结束。其中, (i, j) 为某个像素点; NBD 为当前跟踪边界编号, 初始值为 1; $LABN$ 为上一个边界编号; 光栅扫描是指从左往右, 由上往下, 先扫描完一行, 再移至下一行起始位置继续扫描。

6. 数字和小数点分割

每个数字和小数点的轮廓已经找到, 接下来就很利于分割操作了。首先通过质心查找函数可以找到每个轮廓质心的坐标, 接着对轮廓按质心横坐标进行排序并计算每个轮廓的面积, 然后找到面积最小的(小数点)且记录其索引号(小数点位置), 最后便可以提取每个数字, 完成分割。

7. 数字识别

数字识别是数字仪表识别的重点, 能正确识别数字才可保证识别率。目前常见的字符识别算法主要有模板匹配法[9]、穿线法[10]、特征检测法[11]和神经网络法[12]等。本文采用的是基于注意机制的卷积神经网络算法。

7.1. 传统卷积神经网络

如图 2 所示, 传统的卷积神经网络由输入层、卷积层、池化层、全连接层和输出层构成。卷积神经网络的核心是卷积层, 用卷积核提取特征。卷积过程是卷积核沿着高和宽的方向在图像上移动, 首先卷积核中的权重与其对应位置的参数相乘, 然后将所有相乘结果相加得到一个输出, 当卷积核遍历整个图像后, 最后得到输出特征图。池化层就是对特征图进行压缩, 简化网络的计算复杂度, 提取主要特征。通常形式有最大池化和平均池化两种, 最大池化是提取区域内的最大值, 平均池化是提取区域内的平均

值。全连接层往往用于卷积神经网络结构的最后几层，将多维的特征映射成一个一维向量，对特征进行分类。

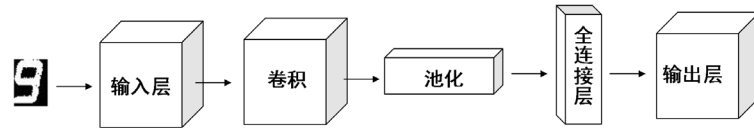


Figure 2. Traditional convolutional neural network structure
图 2. 传统卷积神经网络结构

7.2. 改进的卷积神经网络

传统的卷积神经网络需要大量的数据集进行训练且训练过程中计算复杂度高，参数量过大。针对这一问题，本文提出了改进的卷积神经网络算法，即采用深度可分离卷积并引入注意机制模块[13]，其结构图如图 3 所示。该网络先使用深度可分离卷积提取特征，然后通过通道注意力模块和空间注意力模块来得到通道注意图和空间注意图，相乘得到最后特征图[14] [15] [16] [17]。其中第一层为 1×1 的逐点卷积，主要是用来提升通道数量；第二层为 3×3 的深度可分离卷积层，提取特征并减少了模型参数；第三层为 1×1 逐点卷积层，降低维度；第四层为平均池化层和最大池化层，聚合空间维度信息，为每个通道生成两个特征描述符，第五层为共享多层感知器层，将经过两种池化层下采样处理后的两个特征描述输入到共享的多层感知器中，以生成更具代表性的特征向量；第六层为相加层，使用元素求和的方式来合并从多层感知器输出的特征向量，即生成通道注意力图；第七层为相乘层，逐点卷积层与相加层所产生的特征图相乘来进行特征的自适应学习；第八层为平均池化层和最大池化层，为每个通道生成两个特征描述符；第九层为连接层，对经过两种池化层下采样处理后的两个特征描述进行连接；第十层为标准卷积层，通过 7×7 的卷积运算得到空间注意图；第十一层相乘层，第七层相乘层与第十层标准卷积层所产生的特征图相乘来进行特征的自适应学习，得到最后的特征图。其主体网络如表 1 所示。

Table 1. The main network
表 1. 主体网络

类型	卷积核	输入	输出
卷积 3×3	32	$112 \times 112 \times 3$	$112 \times 112 \times 32$
逐点卷积 1×1	64	$112 \times 112 \times 32$	$112 \times 112 \times 64$
深度卷积 3×3	64	$112 \times 112 \times 64$	$112 \times 112 \times 64$
逐点卷积 1×1	64	$112 \times 112 \times 64$	$112 \times 112 \times 64$
注意机制模块	-	$112 \times 112 \times 64$	$112 \times 112 \times 64$
逐点卷积 1×1	128	$112 \times 112 \times 64$	$112 \times 112 \times 128$
深度卷积 3×3	128	$112 \times 112 \times 128$	$56 \times 56 \times 128$
逐点卷积 1×1	128	$56 \times 56 \times 128$	$56 \times 56 \times 128$
注意机制模块	-	$56 \times 56 \times 128$	$56 \times 56 \times 128$
逐点卷积 1×1	256	$56 \times 56 \times 128$	$56 \times 56 \times 256$
深度卷积 3×3	256	$56 \times 56 \times 256$	$28 \times 28 \times 256$
逐点卷积 1×1	256	$28 \times 28 \times 256$	$28 \times 28 \times 256$
注意机制模块	-	$28 \times 28 \times 256$	$28 \times 28 \times 256$

Continued

逐点卷积 1×1	512	$28 \times 28 \times 256$	$28 \times 28 \times 512$
深度卷积 3×3	512	$28 \times 28 \times 512$	$14 \times 14 \times 512$
逐点卷积 1×1	512	$14 \times 14 \times 512$	$14 \times 14 \times 512$
注意机制模块	-	$14 \times 14 \times 512$	$14 \times 14 \times 512$
逐点卷积 1×1	1024	$14 \times 14 \times 512$	$14 \times 14 \times 1024$
深度卷积 3×3	1024	$14 \times 14 \times 1024$	$7 \times 7 \times 1024$
逐点卷积 1×1	1024	$7 \times 7 \times 1024$	$7 \times 7 \times 1024$
全连接层	-	$7 \times 7 \times 1024$	$1 \times 1 \times 1024$
Softmax	-	$1 \times 1 \times 1024$	10

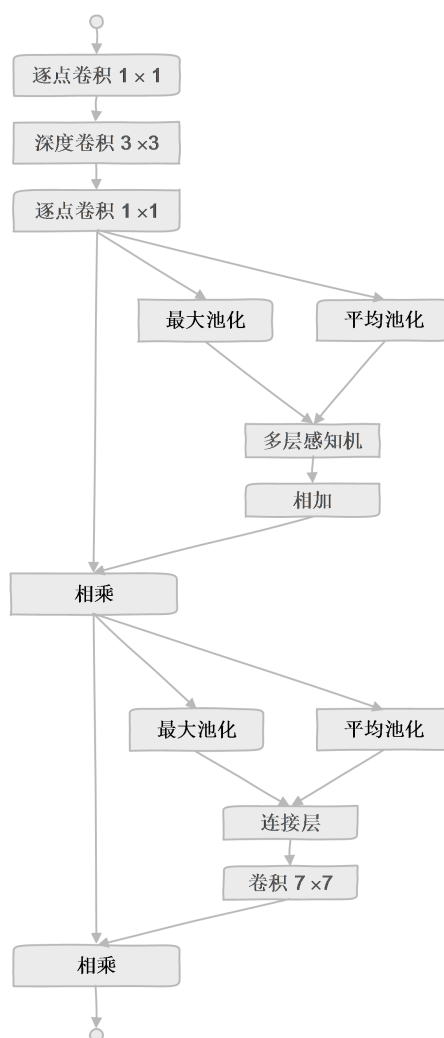


Figure 3. The structure of convolutional neural network based on attention mechanism

图 3. 基于注意机制的卷积神经网络结构

8. 实验

8.1. 数据集

训练卷积神经网络时,首先需要采集数据集并对数据集进行标注和扩增。数字数据集共 1000 张图片包括 10 个分类,分别是 0、1、2、3、4、5、6、7、8、9,每个分类 100 张。部分数据集如图 4 所示。



Figure 4. Partial digital data set
图 4. 部分数字数据集

8.2. 试验及分析

本文的算法设计利用 python 和 opencv 编写完成,实现了数显仪表的图像采集、图像处理和图像识别,其每一步的效果图如图 5 所示。



Figure 5. Experimental process diagram
图 5. 实验过程图

本文一共采集了 200 张数显仪表图像,分别用模板匹配算法、穿线法、神经网络算法和本文的基于注意机制和深度可分离卷积神经网络算法进行测试,测试结果如表 2 所示。

试验结果表明,这些算法对数显仪表图像的识别率都不错,本文算法其他算法相比有显著的提高,识别效果较好。证明了本文所提出的网络模型能更好地提取图像特征,在数显仪表识别率上有着明显的优势。

Table 2. Experimental result
表 2. 试验结果

算法	图像/张	正确率/%
模板匹配法	200	96.2
穿线法	200	96.1
神经网络法	200	97.0
本文算法	200	98.5

9. 结论

本文提出了一种基于传统图像处理和改进的卷积神经网络的数显仪表识别方法。首先通过传统图像处理分割出待识别的数字，再将数字送进卷积网络模型进行识别。改进的网络模型先通过深度可分离卷积减少模型参数，可以大大减少训练时间，接着采用的注意机制模块通过通道注意和空间注意的融合能够有效地提高网络特征提取能力，提高网络的数字识别率，并在试验中证实了该算法得高效性。

基金项目

广东省自然科学基金(2016A030313003)。

参考文献

- [1] 王艳, 谢广苏, 沈晓宇. 一种基于 MSER 和 SWT 的新型车牌检测识别方法研究[J]. 计量学报, 2019, 40(1): 82-90.
- [2] 张海波, 段会川, 张曙光, 等. 一种数字仪表显示值快速识别方法[J]. 计算机工程与应用, 2005, 41(4): 223-226.
- [3] 王臣, 严军, 朱静, 等. 基于图像的数字仪表数字轮定位分割算法的研究[J]. 电子测量技术, 2017, 40(9): 132-135.
- [4] 申小阳, 唐铁峻, 姜柏军, 等. 仪表的数字字符识别系统[J]. 仪表技术与传感器, 2005(6): 55-57.
- [5] 陆靖滨, 许丽. 基于自适应特征提取的数显仪表识别系统[J]. 现代电子技术, 2017, 40(24): 147-150.
- [6] 蔡梦倩, 张蕾, 王炎, 莫娟. 基于全卷积网络的数字仪表字符识别方法[J]. 现代计算机(专业版), 2018(2): 38-43.
- [7] Suzuki, S. (1985) Topological Structural Analysis of Digitized Binary Images by Border Following. *Computer Vision, Graphics, and Image Processing*, **30**, 32-46. [https://doi.org/10.1016/0734-189X\(85\)90016-7](https://doi.org/10.1016/0734-189X(85)90016-7)
- [8] 丁双, 任国营, 张福民, 范乃胤. 改进的穿线法的卡尺图像识别[J]. 计量学报, 2019, 40(5): 765-769.
- [9] 邓清男, 石晓龙. 变电站室内数显仪表的读数识别[J]. 工业仪表与自动化装置, 2018(2): 86-89.
- [10] 卓磊, 周律, 杨丽红. 基于穿线法的计算器数字识别[J]. 软件工程, 2018, 21(12): 1-3.
- [11] 陈刚, 胡子峰, 郑超. 基于特征检测的数字仪表数码快速识别算法[J]. 中国测试, 2019, 45(4): 146-150.
- [12] 熊勋, 陈新度, 吴磊, 林旭华. 基于卷积记忆神经网络的数字表盘读数识别[J]. 组合机床与自动化加工技术, 2019(7): 72-75.
- [13] Misra, D., Nalamada, T., Arasanipalai, A.U., et al. (2020) Rotate to Attend: Convolutional Triplet Attention Module.
- [14] Chen, L., Zhang, H.W., Xiao, J., Nie, L.Q., Shao, J., Liu, W. and Chua, T.-S. (2017) Sca-cnn: Spatial and Channel-Wise Attention in Convolutional Networks for Image Captioning. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 5659-5667. <https://doi.org/10.1109/CVPR.2017.667>
- [15] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-22 June 2018, 7132-7141. <https://doi.org/10.1109/CVPR.2018.00745>
- [16] Liu, J.-J., Hou, Q., Cheng, M.M., Wang, C.H. and Feng, J.S. (2020) Improving Convolutional Networks with Self-Calibrated Convolutions. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 13-19 June 2020, 10096-10105.
- [17] Woo, S., Park, J., Lee, J.-Y. and Kweon, I.S. (2018) Cbam: Convolutional Block Attention Module. *The European Conference on Computer Vision (ECCV)*, Munich, 8-14 September 2018, 3-19.