

# Reflection and Rotation Symmetry Detection via Equivariant Learning

## 摘要

对称性广泛存在于视觉图像，利用图像的对称性对图片进行处理在计算机视觉存在广泛的应用。检测对称性的固有挑战源于对称轴方向的任意性，反射对称根据特定方向的对称轴镜像对称。因此，从图像中发现这种对称模式受益于等变特征表示，它会随着图像的反射和旋转而一致地变化。在本论文中使用了一个用于对称检测的群等变卷积网络，该网络利用了关于反射和旋转的二面体群的等变特征映射，从而能更好的提取等变特征。网络采用二面体群的等变层进行构建，并训练输出反射轴得分图。在改进方面，本实验通过使用阈值过滤和霍夫变换对反射对称轴得分图进行直线检测，最后根据投票得到图像的主对称轴。在实验中，选取三个不同难度等级的图像作为 MYTEST 数据集，并对对称检测结果进行可视化。此外，本课题使用角度和中心偏移作为对称准确率评估标准，并以此统计分析算法的准确率。

**关键词：**图像对称性检测；镜像对称；反射对称； $E(2)$ -群等变卷积

## 1 引言

从分子到星系，从自然到人造环境，对称无处不在。在图像内容的各种特性中，对称性是一个具有规律性的明显特征。本文的重点是反射对称检测，目标是找到在反射下保持不变的反射对称轴。

对称性数学概念的简单性鼓励了早期的方法来寻找满足对称性预定义约束的关键点对，其利用手工制作的局部特征描述符来检测稀疏对称模式。最近，卷积神经网络（CNNs）已成功应用于检测反射对称性，并通过从数据中学习得分图回归 [5] 或对称匹配 [15] 超过了以前的方法。检测对称模式的主要挑战在于对称性以任意方向表现出来，并且感知该模式需要基于该方向进行分析。反射对称性相对于具有特定方向的轴反射自身。因此，大多数对称检测方法涉及在对称模式的候选方向的空间上进行搜索，并且还开发相对于旋转和反射不变或等变的鲁棒表示。早期的方法通过提取定向关键点和执行定向归一化来利用等变表示 [5, 14, 16, 17, 20]。虽然这项技术已被证明对基于浅梯度的特征有效，但它不能应用于标准神经网络的深层特征图，因为旋转和反射会导致不可预测的表示变化。

为了应对这一挑战，我们建议学习一种用于反射和旋转对称检测的群等变卷积神经网络，称为 EquiSym。与标准神经网络不同，等变网络诱导了关于几何变换的可预测和结构保持表示，例如旋转或反射，这非常适合用于对称检测。为了检测不同方向上的一致对称模式，我

们构建了一个二面体等变卷积网络 [21]，该网络被设计为对一组反射和旋转是端到端等变的。网络有效地学会输出用于反射对称的反射轴的得分图或用于旋转对称的旋转中心的得分图。

## 2 相关工作

### 2.1 稀疏预测

在检测对称性时，一致需要仿射不变或等变特征描述符，因为匹配局部描述符是最常见的解决方案。Loy 和 Eklundh [14] 使用 SIFT [13] 描述符，并通过其主导方向对每个描述符进行归一化。Cho 和 Lee [2] 还使用 SIFT [13] 来匹配特征对，并通过发现附近匹配的聚类来检测对称性。轮廓和边缘特征 [5, 16, 17, 19, 20] 也可用于确定对称对象的边界。Lee 和 Liu [9] 提出通过校正偏斜图案来解决仿射偏斜旋转对称群检测。在本文中，我们用我们提出的数据集通过数据驱动的方法来处理这项任务。此外，我们使用二面体群来解释对称性，就像在许多对称性检测文献中一样 [8, 9]。

### 2.2 密集预测

最近提出的方法 [4, 5, 15, 18] 预测逐像素对称性得分。Fukushima 和 Kikuchi [4] 建立了一个神经网络，从图像中提取边缘并检测反射对称性。Tsogkas 等人 [18] 使用每个像素的直方图、颜色和纹理构建一组特征，并在训练模型时采用多实例学习。Gnutti 等人 [6] 采取两个阶段，即使用逐片相关计算每个像素的对称性得分，并使用梯度方向和大小验证所获得的候选轴。Funk 和 Liu [5] 是第一个采用深层细胞神经网络来检测反射和旋转对称性的人。Seo 等人 [15] 提出了一种极自相似描述符，以获得更好的旋转和反射不变性。一种专门设计的极性匹配卷积 (PMC) 执行逐区域特征匹配来计算对称性得分，但该模型严重依赖于卷积神经网络。为了发现旋转和反射的几何变换的一致对称模式，我们在对称检测模型中部署了群等变神经网络。

## 3 本文方法

### 3.1 本文方法概述

图 1 展示了本文提出的 EquiSym 对称检测网络。在输入图像后，通过一个群等变编码器  $Enc$  和群等变解码器  $Dec^{ref}$  预测反射对称轴，得到中间预测值  $S^{ref}$ 。反射对称的辅助任务是设置反射轴的方向（在本论文中设置为个  $N = 8$  方向）。辅助任务将前景类的  $N$  个通道合并为  $P^{ref}$ ，并与分数  $S^{ref}$  进行拼接。最后将拼接的结果放入群等变  $1 \times 1$  卷积，进而得到反射对称轴的最终预测分数  $Y^{ref}$ 。

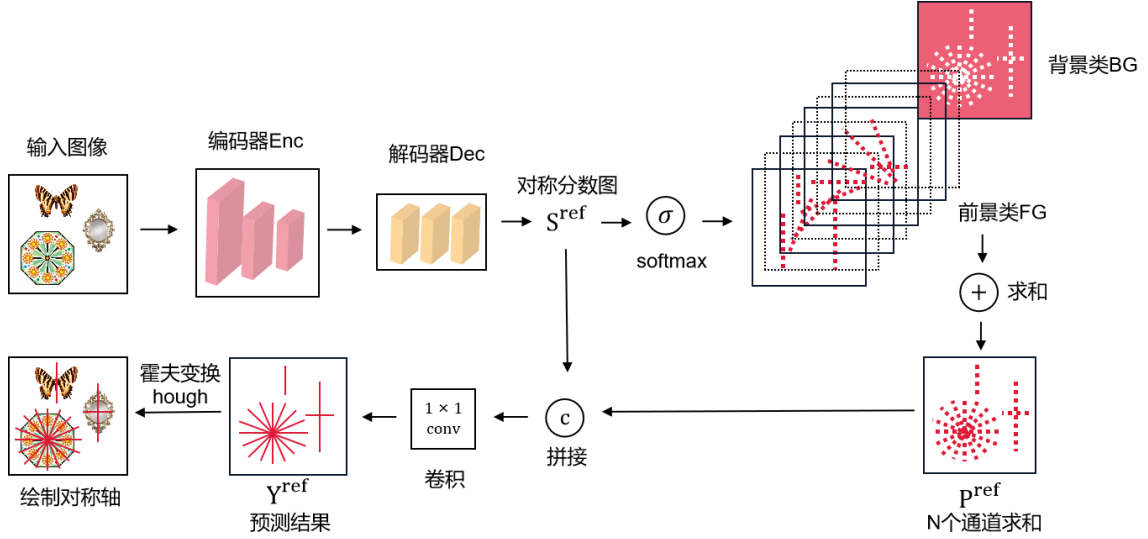


图 1. EquiSym 对称检测网络示意图

### 3.2 反射等变模块

由于本算法的目标是检测图像的反射对称轴，所以首先使用了一个二面体群  $D_N$  的  $E(2)$  等变卷积神经网络，它具有  $N$  个离散旋转 (角度为  $\frac{2\pi}{N}$  的倍数) 和镜像对称。编码器  $Enc$  由一个等变 ResNet [7] 和一个空洞空间金字塔池化 (ASPP [1], Atrous Spatial Pyramid Pooling) 模块组成。解码器  $Dec$  是一个三层卷积模块。编码器和解码器中所有的卷积层都被  $E(2)$ -等变卷积层所取代。在正向计算过程中，将特征字段转换为二面体群  $D_N$  的预定义字段。对于检测结果  $S^{ref}$ ，解码器  $Dec^{ref}$  的特征字段被转换回标量字段。

### 3.3 辅助分类

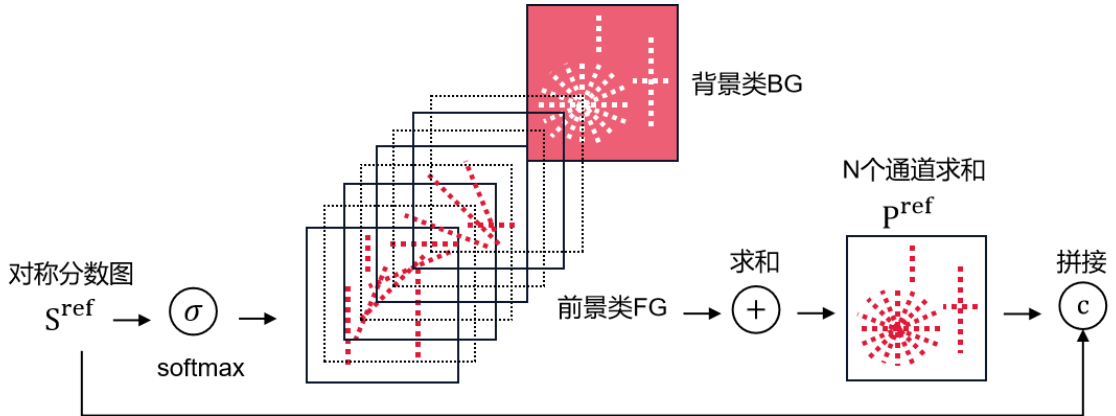


图 2. 辅助分类示意图

本算法并不是直接回归对称分数图，而是执行相关的子任务，从而得出最终的预测。辅助分类任务是对反射对称轴的方向或者角度的像素级分类。如图2为辅助分类示意图。为简单起见，本算法将方向分配成  $N^{ref}$  个，这  $N^{ref}$  个方向将  $180^\circ$  平分。然后将  $groundtruth$  中对称轴的方向  $S_{gt}^{ref}$  量化为  $one-hot$ 。同时，本算法为既不是轴也不是中心的像素添加背景类，背景类全被编码为 0。因此，分类器预测  $N^{ref} + 1$  个通道的分数。分类逻辑  $S \in R^{H \times W \times (N+1)}$

由式1得到。

$$S = Dec(Enc(I)) \quad (1)$$

### 3.4 对称性检测

对应方向对称轴的预测得分图出现在估计方向  $S^{ref}$  的个前景通道中，而背景通道中包含背景像素。本算法将前景类  $P \in R^{H \times W \times 1}$  (式2) 和中间预测  $S \in R^{H \times W \times N+1}$  (式1) 的和相加，以计算最终预测  $Y \in R^{H \times W}$ ，如式3所示， $\parallel$  表示沿着最终通道维度的拼接操作。

$$P_{h,w} = \sum_{k=1}^N \frac{\exp(S_{h,w,k})}{\sum_c \exp(S_{h,w,c})} \quad (2)$$

$$Y = conv_G([p \parallel S]) \quad (3)$$

### 3.5 损失函数

为了训练对称检测网络，本算法优化了用于定位和分类的两个损失项的组合。本算法采用焦点损失 [10] 作为反射得分图的定位损失  $L_{loc}$  (式4)， $Y$  表示最终预测对称得分图， $Y_{gt}$  为对称轴真实值。中间预测的分类损失  $L_{cls}$  (式5) 使用交叉熵损失计算， $S$  表示中间预测得分图， $S_{gt}$  为分类真实值。最终的目标损失  $L$  (式6) 为定位损失  $L_{loc}$  和分类损失  $L_{cls}$  求和。为了缓解类不平衡的问题，使用  $\omega$  对背景类损失  $L_{cls}$  进行加权计算。焦点损失缓解了  $L_{loc}$  的类不平衡问题。

$$L_{loc} = L_{focal}(Y, Y_{gt}) \quad (4)$$

$$L_{cls} = L_{ce}(S, S_{gt}) \quad (5)$$

$$L = L_{loc} + L_{cls} \quad (6)$$

## 4 复现细节

### 4.1 与已有开源代码对比

本文复现了作者所提出的基于等变学习的对称检测网络 EquiSym 模型，并且对网络的输出端进行了修改。原始网络的输出只有对对称轴的像素级别的预测图像，并没有实现将对称轴画出来。在作者的基础上，本论文检测对称轴预测图像中的直线，通过阈值确认对称轴并且使用霍夫变换检测直线，最终将对称轴绘制在图像上。同时，本实验验证了作者所提出模型的有效性。

本文复现工作基于作者发布于 GitHub 的开源代码：<https://github.com/ahyunSeo/EquiSym>

### 4.2 创新点

通过训练好的等变神经网络可以得到对称检测的最终分数图  $Y$ 。该分数图为灰度图，值越大的地方是对称轴的可能性越高。本论文通过将得分图  $Y$  转化为热力图可以更加直观的看出对称轴所在位置。在热力图中颜色越红的地方则为对称轴的可能性越大。在分数图  $Y$  中可

以检测出多条对称轴，本实验首先通过阈值确认该点是否属于对称轴，然后使用霍夫变换根据投票机制检测直线，最终得到检测得分图中的直线，并在原图像上绘制出主对称轴。

如图3所示，从左往右，第 1 张为原图，第 2 张为通过等变神经网络得到的最终得分图，第 3 张为由得分图得到的热力图，第 4 张为通过得分图经过霍夫变换检测到的对称轴图。

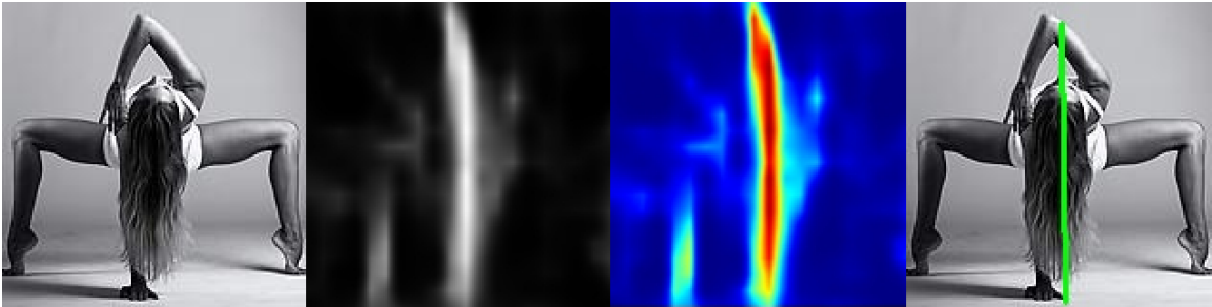


图 3. 可视化结果示意图

4.3 实验可视化界面分析

如图4所示，选择图片后，会在窗口第 1 行第 1 列显示原图，第 2 列显示真实对称轴，第 3 列显示角度误差和中心偏移误差。第 2 行第 1 列显示得分图，第 2 列显示热力图，第 3 列是预测对称轴结果图。

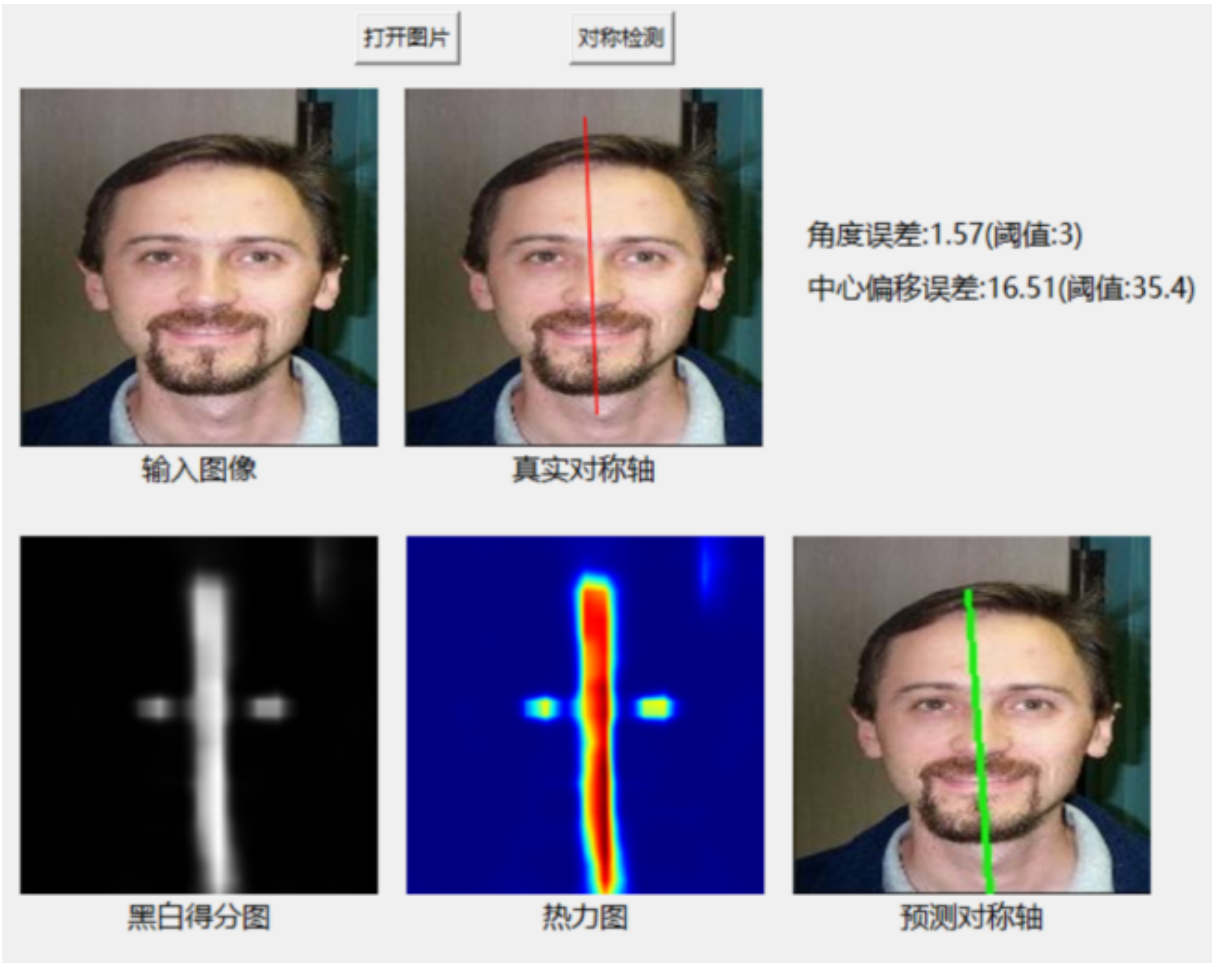


图 4. 可视化结果示意图



## 5 实验结果分析

本实验中采用了 CVPR 反射对称竞赛中的数据集中的图像, NYU [3]、SDRW [12] 图像、COCO [11] 对称数据集中的图像, 以及部分来自网络上下载了一些对称图像。这些数据集的对称检测会有不同的难易程度, 将其分成具有绝对对称的简单图像、具有不复杂背景的中等图像和具有干扰的困难图像三个等级。每个等级分别有各二十张图像, 共六十张的数据集 MYTEST, 分类依据如表1所示。

等级	标准
简单 (EASY)	具有绝对对称性的几何图形、简单图形、纯色背景图
中等 (NORMAL)	有背景干扰的自然物体, 背景不复杂的现实生活图像
困难 (HARD)	有噪声、形变的自然图像, 背景复杂的现实生活图像

表 1. MYTEST 数据集分类标准

评估算法的标准主要是对称轴检测的准确率。准确率从两个方面考虑: 中心偏移误差  $E_{co}$  与角度误差  $E_{ang}$ 。设线段  $GT$  为真实对称轴, 线段  $R$  为预测对称轴。中心偏移误差使用检测结果的线段与真正的对称轴线段中点间的距离  $d$  衡量, 角度误差使用两个线段之间的夹角来衡量。若  $d$  与  $\theta$  均小于设置的值则检测正确。CVPR2013 的评估标准为:  $Ang = 10^\circ$ ,  $Len = 20\% \times \min(length(GT), length(R))$ 。(其中  $length(GT)$  和  $length(R)$  分别表示  $GT$  轴和  $R$  轴的长度)。本实验中参考了 CVPR2013 的评估标准 [12], 把中心偏移阈值设置为与 CVPR2013 标准相同, 但是角度阈值设置为  $3^\circ$ 。若满足  $d < Len$  且  $\theta < Ang$ , 则检测出的对称轴  $R$  为真。

根据评估标准, 对本算法的对称检测结果进行准确率的评估分析。对 MYTEST 数据集中三个难度等级的图像, 分别计算了其中心偏移误差和角度误差, 并根据这两者的阈值统计预测正确的图片数量和相应的预测准确率。统计结果分别如表2和表3所示。

等级	$E_{ang} \leq Ang$	$E_{co} \leq Len$	$E_{ang} \leq Ang$ 且 $E_{co} \leq Len$
EASY	19	18	18
NORMAL	18	17	15
HARD	17	16	15

表 2. Caption

等级	角度准确率	中心偏移准确率	总准确率
EASY	95%	90%	90%
NORMAL	90%	85%	75%
HARD	85%	80%	75%
TOTAL	90%	85%	80%

表 3. Caption

将这三个难度的部分图像检测结果进行可视化，得到的图像如图5（简单）、6（中等）、7（困难）所示。

对于这类具有绝对对称性的 EASY 图形，得到的评分图也基本是对称的，并且在主对称轴的两边也呈现明显的对称性。从得到的结果可以看出，通过学习得到的在该难度图像下的对称轴检测结果基本上重合于真正的对称轴，只有个别有极小的误差，在这一级别的数据集上表现良好并且比较准确。

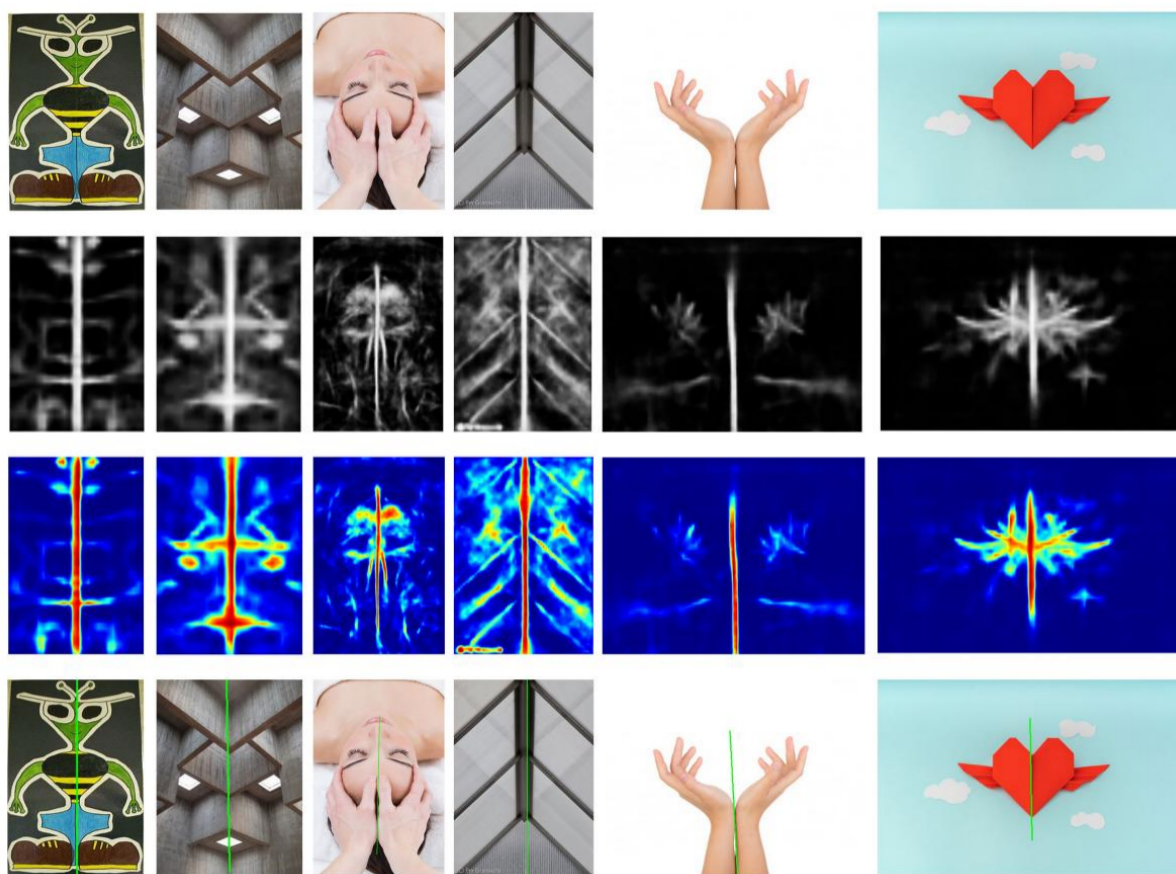


图 5. EASY 数据集上可视化结果

NORMAL 难度的图像选取了有背景干扰的自然物体，背景不复杂的现实生活图像。由于背景的干扰，这在一定程度上会影响了检测的效果。

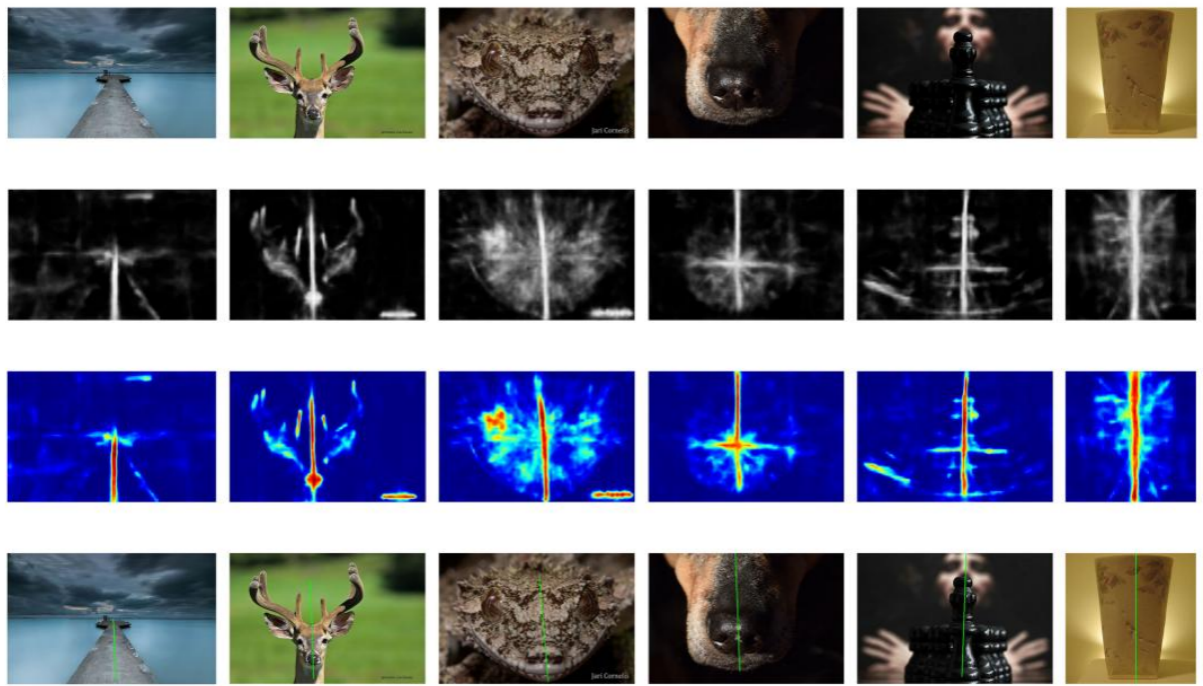


图 6. NORMAL 数据集上可视化结果

HARD 难度图像由于背景杂乱、噪声、阴暗程度不同等原因，不具有绝对的反射对称性，这些干扰都会使得对称检测的难度提高。



图 7. HARD 数据集上可视化结果



## 6 总结与展望

本文构建了一个二面体等变卷积网络 EquiSym 从而使得网络能够检测不同方向上的一致对称模式。但是在检测过程中,还是会出现无法检测出部分图像不同旋转角度的正确对称轴。这有可能是训练的参数还不够好,或者是网络本身具有的问题,需要做进一步的改进。对称图像往往还有各种复杂的情形,如具有透视、光线不一和背景复杂的图像,如何应对这些情况从而改进等变学习算法的准确率,提高算法的鲁棒性,也是未来的研究方向之一。

实验中仅根据霍夫变换检测出对称性度量最高的直线作为主对称轴。但是论文算法是可以检测出多条对称轴的,可以改进算法设定阈值,将得分图中对称轴度量超过阈值的对称轴显示出来。

## 参考文献

- [1] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *CoRR*, abs/1706.05587, 2017.
- [2] Minsu Cho and Kyoung Mu Lee. Bilateral symmetry detection via symmetry-growing. In Andrea Cavallaro, Simon Prince, and Daniel C. Alexander, editors, *British Machine Vision Conference, BMVC 2009, London, UK, September 7-10, 2009. Proceedings*, pages 1–11. British Machine Vision Association, 2009.
- [3] Marcelo Cicconet, Vighnesh Birodkar, Mads Lund, Michael Werman, and Davi Geiger. A convolutional approach to reflection symmetry. *Pattern Recognit. Lett.*, 95:44–50, 2017.
- [4] Kunihiro Fukushima and Masayuki Kikuchi. Symmetry axis extraction by a neural network. *Neurocomputing*, 69(16-18):1827–1836, 2006.
- [5] Christopher Funk and Yanxi Liu. Beyond planar symmetry: Modeling human perception of reflection and rotation symmetries in the wild. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 793–803, 2017.
- [6] Alessandro Gnutti, Fabrizio Guerrini, and Riccardo Leonardi. Combining appearance and gradient information for image symmetry detection. *IEEE Trans. Image Process.*, 30:5708–5723, 2021.
- [7] Jiaming Han, Jian Ding, Nan Xue, and Gui-Song Xia. Redet: A rotation-equivariant detector for aerial object detection. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, pages 2786–2795. Computer Vision Foundation / IEEE, 2021.
- [8] Seungkyu Lee, Robert T. Collins, and Yanxi Liu. Rotation symmetry group detection via frequency analysis of frieze-expansions. In *2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24-26 June 2008, Anchorage, Alaska, USA*. IEEE Computer Society, 2008.

- [9] Seungkyu Lee and Yanxi Liu. Skewed rotation symmetry group detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(9):1659–1672, 2010.
- [10] Tsung-Yi Lin, Priya Goyal, Ross B. Girshick, Kaiming He, and Piotr Dollár. Focal loss for dense object detection. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 2999–3007. IEEE Computer Society, 2017.
- [11] Tsung-Yi Lin, Michael Maire, Serge J. Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft COCO: common objects in context. In David J. Fleet, Tomás Pajdla, Bernt Schiele, and Tinne Tuytelaars, editors, *Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V*, volume 8693 of *Lecture Notes in Computer Science*, pages 740–755. Springer, 2014.
- [12] Jingchen Liu, George M. Slota, Gang Zheng, Zhaohui Wu, Minwoo Park, Seungkyu Lee, Ingmar Rauschert, and Yanxi Liu. Symmetry detection from realworld images competition 2013: Summary and results. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR Workshops 2013, Portland, OR, USA, June 23-28, 2013*, pages 200–205. IEEE Computer Society, 2013.
- [13] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.*, 60(2):91–110, 2004.
- [14] Gareth Loy and Jan-Olof Eklundh. Detecting symmetry and symmetric constellations of features. In Aleš Leonardis, Horst Bischof, and Axel Pinz, editors, *Computer Vision – ECCV 2006*, pages 508–521, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.
- [15] Ahyun Seo, Woohyeon Shim, and Minsu Cho. Learning to discover reflection symmetry via polar matching convolution. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 1265–1274, 2021.
- [16] Dinggang Shen, Horace H.S. Ip, and Eam Khwang Teoh. Robust detection of skewed symmetries by combining local and semi-local affine invariants. *Pattern Recognition*, 34(7):1417–1428, 2001.
- [17] V. Shiv Naga Prasad and L.S. Davis. Detecting rotational symmetries. In *Tenth IEEE International Conference on Computer Vision (ICCV’05) Volume 1*, volume 2, pages 954–961 Vol. 2, 2005.
- [18] Stavros Tsogkas and Iasonas Kokkinos. Learning-based symmetry detection in natural images. In Andrew W. Fitzgibbon, Svetlana Lazebnik, Pietro Perona, Yoichi Sato, and Cordelia Schmid, editors, *Computer Vision - ECCV 2012 - 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part VII*, volume 7578 of *Lecture Notes in Computer Science*, pages 41–54. Springer, 2012.

- [19] Zhaozhong Wang, Lianrui Fu, and Y. F. Li. Unified detection of skewed rotation, reflection and translation symmetries from affine invariant contour features. *Pattern Recognit.*, 47(4):1764–1776, 2014.
- [20] Zhaozhong Wang, Zesheng Tang, and Xiao Zhang. Reflection symmetry detection using locally affine invariant edge correspondence. *IEEE Transactions on Image Processing*, 24(4):1297–1301, 2015.
- [21] Maurice Weiler and Gabriele Cesa. General  $e(2)$ -equivariant steerable cnns. *CoRR*, abs/1911.08251, 2019.