Axis-level Symmetry Detection with Group-Equivariant Representation

Anonymous ICCV submission

Paper ID 16

Abstract

Symmetry is a fundamental concept that has been studied extensively; however, its detection in complex scenes remains challenging in computer vision. Recent heatmap-based methods identify potential regions of symmetry axes but lack precision for individual axis. In this work, we introduce a novel framework for axis-level detection of the most common symmetry types—reflection and rotation—representing them as explicit geometric primitives i.e., *lines and points. We formulate a dihedral group-equivariant dual-branch architecture, where each branch exploits the properties of dihedral group-equivariant features in a novel, specialized manner for each symmetry type. Specifically, for reflection symmetry, we propose* ***orientational anchors*** *aligned with group components to enable orientation-specific detection, and* ***reflectional matching*** *that computes similarity between patterns and their mirrored counterparts across potential reflection axes. For rotational symmetry, we propose* ***rotational matching*** *that computes the similarity between patterns at fixed angular intervals. Extensive experiments demonstrate that our method significantly outperforms state-of-the-art methods.*

1. Introduction

Symmetry is a fundamental concept observed across natural and artificial environments [51, 56], appearing at multiple scales and orientations [17, 39]. While humans easily recognize symmetry [52], it remains a challenge for computer vision. This work focuses on detecting reflection and rotation symmetries in complex 2D scenes [35]. Robust symmetry detection requires precise localization of reflection and rotation axes, along with accurate determination of additional properties; reflection symmetry involves estimating axis length and orientation [60], while rotation symmetry requires classifying the correct fold. These challenges are amplified by real-world complexities such as noise, occlusion, and geometric distortions.

Symmetry detection has evolved from classical to deep learning approaches. Traditional methods use descriptor matching for reflection symmetry [3, 27, 36, 37], and frequency domain analysis [25, 30] or gradient vector flow [40] for rotation symmetry. Neural networks provide advancements to the field, from early symmetry-aware models [14, 50] to CNN-based pixel prediction [15, 48], and recently to self-similarity networks and group-equivariant architectures [43, 44].

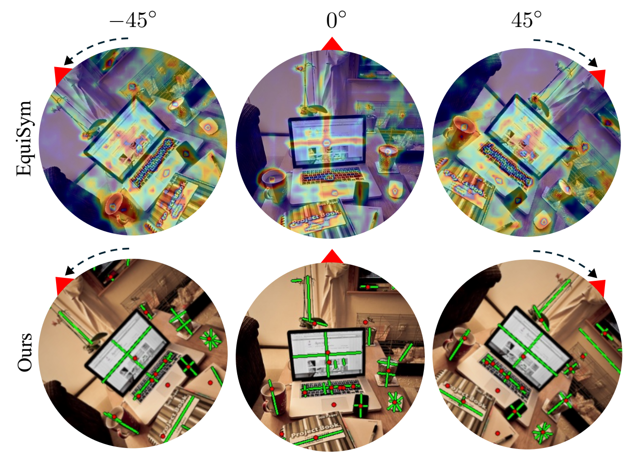

Figure 1. **Comparison of the heatmap-based method [44] and our axis-level approach on rotated inputs.** The red triangle indicates the rotated orientation of input image. Our proposed axis-level symmetry detection method captures reflection (green lines) and rotation (red points) axes as precise geometric entities and demonstrates superior robustness to rotation compared to the heatmap-based method.

Despite these advancements, neural network-based approaches face two key limitations. First, most treat symmetry detection as a per-pixel heatmap prediction problem [15, 43, 44], which makes it difficult to recover the precise geometric parameters of symmetry axes. Second, they lack explicit integration of symmetric structures in feature representations [43], or fail to consider explicit matching modules [44], leading to inconsistent results under arbitrary rotations and reflections of the input image.

To address these limitations, we propose an *axis-level* symmetry detection approach that is equivariant to the dihedral group. Reflection and rotation symmetry axes are modeled explicitly as geometric primitives, *i.e.*, lines and points. We leverage features that are equivariant under the dihedral group D_N (the group of planar rotations and reflections) and build specialized modules for each symmetry type, ensuring that the symmetry predictions transform predictably with rotations or reflections of the input.

For reflection symmetry, we introduce **orientational anchors** aligned with the discrete orientations of D_N, enabling the network to detect reflection axes in an orientation-specific manner while remaining equivariant to image rotations and reflections. We also include a **reflectional matching** module that explicitly compares patterns with their mirror-image across candidate axes. By design, this reflectional matching module preserves equivariance to D_N and is invariant to reflections of the input (intuitively, rotating an image rotates the symmetry response accordingly, while mirroring preserves the symmetry response).

For rotation symmetry, we present a **rotational matching** module that compares features with rotated versions of themselves at fixed angles. This module is constructed to be completely invariant to dihedral group, ensuring consistent detection of rotational symmetry centers regardless of image rotations or reflections. Fig. 1 provides an overview, illustrating that our method produces consistent reflection and rotation axis detections even when the input image is rotated, in contrast to standard heatmap-based methods. Experiments on real-world datasets demonstrate that our approach consistently outperforms pixel-level methods.

The contributions of this paper include:

- We propose a novel *axis-level* symmetry detection network for reflection and rotation symmetry, leveraging representations that are equivariant to the dihedral group D_N.
- We introduce an *orientational anchor expansion* mechanism for orientation-specific analysis, integrating the group's rotation dimension into the detection process.
- We develop an *equivariant reflectional matching* module and an *invariant rotational matching* module to explicitly incorporate symmetry-consistent feature comparisons for reflection and rotation, respectively.
- We validate our approach on real-world datasets, where it achieves superior performance compared to existing methods.

2. Related work

Symmetry detection. Early reflection symmetry detection used keypoint matching [3, 37] with SIFT descriptors [36], while contour [45, 53] and gradient-based [18, 47] methods extracted symmetry structures. Randomized approaches [4] aligned patterns via cross-correlation, while Hough voting [23], local affine frames [10], and RANSAC [46] refined axis extraction for planar surfaces. For rotation symmetry, early methods identified periodic signals in spatial [34] and frequency domains [25, 30], leveraging spectral density and angular correlation. SIFT-based techniques [36, 37] normalized orientation for rotation detection, while GVF [40] and polar domain representations [1] improved boundary detection. Rectification methods [29] further addressed affine distortions.

Deep learning advanced symmetry detection from early feature extraction [14, 50] to CNN-based heatmaps [15, 43, 48], but focused on dense predictions. Recent optimization-based [24] and neural 3D symmetry reconstruction [31, 61] methods predict instance-level reflection symmetry but are limited to isolated objects without background context. Feature matching remains underexplored—PMCNet [43] introduced polar matching but lacked explicit symmetry integration, while group-equivariant [16, 44] and invariant [12] architectures improved robustness but focused on appearance features. To address these limitations, we redefine symmetry detection as an instance-level task, modeling symmetry axes as geometric entities and integrating equivariant matching for symmetry-aware feature comparisons.

Equivariant neural networks. Convolutional neural networks (CNNs) provide translation equivariance but lack rotation and reflection equivariance, limiting their effectiveness in symmetry-aware tasks. Group-equivariant CNNs [5, 8, 16] introduced group convolutions to address this, with advancements in circular harmonics [57], vector fields [38], and hexagonal lattices [21]. Later work extended equivariance to 3D data [55], intertwiner spaces [9], and homogeneous spaces [6, 7, 54]. Equivariant models have been applied to aerial object detection [19] and symmetry detection [44], with recent works improving keypoint descriptions [2] and enforcing group-equivariant constraints for denoising [49]. Beyond 2D, 3D-equivariant architectures address pose estimation [28] and leverage spherical harmonics for 3D rotation-equivariant encoding [58]. Our approach builds upon dihedral-group equivariant networks, with specialized matching for enhanced symmetry detection.

3. Background

Group. A group is a mathematical structure with a set and an operation satisfying closure, associativity, identity, and invertibility [42]. Groups describe symmetries: transformations like rotations and reflections that preserve an object's structure. Our work is built upon two common discrete groups in neural networks: the cyclic group and the dihedral group. The cyclic group C_N represents discrete rotations $\{r^0, \dots, r^{N-1}\}$, with the group law $r^i r^j = r^{(i+j) \bmod N}$. The dihedral group D_N, relevant to our work, includes both rotations and reflections:

$$\mathrm{D}_N = \{r^0, r^1, \dots, r^{N-1}, b, br^1, \dots, br^{N-1}\}, \quad (1)$$

where r and b are generators of the dihedral group corresponding to rotation and reflection, respectively, satisfying

$b^2 = e$ and $r^n b = br^{-n}$, with e as the identity.

Equivariance. A function $f : \mathcal{X} \to \mathcal{Y}$ is equivariant if it commutes with a group action. Formally, for linear group representations $\sigma_1 : G \to \mathrm{GL}(\mathcal{X})$ and $\sigma_2 : G \to \mathrm{GL}(\mathcal{Y})$, equivariance is defined as:

$$f(\sigma_1(g) \cdot x) = \sigma_2(g) \cdot f(x), \quad \forall g \in G,\, x \in \mathcal{X}. \tag{2}$$

In neural networks, equivariance ensures that transformations in the input induce predictable transformations in the output, preserving data symmetries.

Group representation. A group representation maps each element of a group G to a linear transformation in a vector space [8, 55]. The regular representation [5], a standard way to express discrete groups, maps group elements to permutation matrices on $\mathbb{R}^{|G|\times|G|}$, encoding the group structure through basis vector permutations. For a finite group $G = \{g_1, \dots, g_N\}$, the regular representation $\sigma^G_{\mathrm{reg}}(g)$ is:

$$\sigma^G_{\mathrm{reg}}(g) = [\mathbf{e}_{gg_1}, \dots, \mathbf{e}_{gg_N}], \tag{3}$$

where $\mathbf{e}_{g_i} \in \mathbb{R}^{|G|}$ is a standard basis vector. The identity element e corresponds to the identity matrix: $\sigma^G_{\mathrm{reg}}(e) = \mathrm{I}_{|G|}$. For the cyclic group C_N, the regular representation of r^n is a $N \times N$ cyclic permutation matrix:

$$\sigma^{\mathrm{C}_N}_{\mathrm{reg}}(r^n) = [\mathbf{e}_{r^n}, \mathbf{e}_{r^{(n+1) \bmod N}}, \dots, \mathbf{e}_{r^{(n+N-1) \bmod N}}]. \tag{4}$$

For the dihedral group D_N, the regular representation of br^n permutes both for rotations and reflections:

$$\begin{aligned}\sigma^{\mathrm{D}_N}_{\mathrm{reg}}(br^n) =& [\mathbf{e}_{br^n}, \mathbf{e}_{br^{(n+1) \bmod N}}, \dots, \mathbf{e}_{br^{(n+N-1) \bmod N}},\\ & \mathbf{e}_{r^n}, \mathbf{e}_{r^{(n+1) \bmod N}}, \dots, \mathbf{e}_{r^{(n+N-1) \bmod N}}]. \end{aligned} \tag{5}$$

This captures the structure of D_N in a $2N \times 2N$ matrix. A detailed explanation of group-equivariant representations and corresponding network architectures is provided in the Supplementary Material.

4. Proposed method

In this section, we introduce a D_N-equivariant network for axis-level symmetry detection, modeling reflection axes as line segments and rotation axes as points. The network uses a dihedral group-equivariant backbone [5] to extract features, and then processes these features with two branches: one branch predicts the midpoint, orientation, and length of reflection axes, and the other predicts the location and fold class of rotation symmetry centers (Sec. 4.1). To handle multiple orientations, we introduce orientational anchor expansion, aligning feature channels with the discrete orientations of the dihedral group (Sec. 4.2). We also present reflectional matching to capture symmetry across reflection axes (Sec. 4.3) to compare features with their mirrored versions, and a rotational matching module (Sec. 4.4) to compare the same features at different rotation angles, while preserving dihedral group-equivariance. Fig. 2 illustrates the overall pipeline of our approach.

4.1. Axis-level symmetry detection

Existing neural network-based methods detect symmetry in 2D scenes using pixel-level heatmaps [15, 43, 44], methods detect symmetry by predicting dense pixel-wise heatmaps [15, 43, 44], which makes it hard to recover exact axis parameters. Recent approaches represent reflection axes as lines [24, 31] but are limited to isolated 2D or 3D objects without backgrounds. To address these limitations, we present an axis-level symmetry detection network that accurately represents reflection and rotation axes in complex real-world scenes with multiple instances.

Feature extraction. Given an input image $\mathbf{I}$, we employ a D_N-equivariant backbone network [5, 20] to extract the base feature map $\mathbf{F} \in \mathbb{R}^{H\times W\times\mathcal{C}|\mathrm{D}_N|}$. Here, H and W denote the spatial dimensions, $|\mathrm{D}_N| = 2N$ represents the number of dihedral group elements(combining N rotations and their reflections), and $\mathcal{C}$ is the number of channels per group element. The extracted base feature $\mathbf{F}$ is then fed into the symmetry detection branches.

Reflection symmetry detection. For axis-level reflection symmetry detection, we model reflection axes using the center-angle-length representation [60]. Unlike the endpoint representation [22, 59, 62], this approach inherently handles rotation and reflection equivariance via orientation parameterization. The reflection branch $\mathcal{B}_{\mathrm{ref}}$ processes the base feature map to predict the reflection axes components:

$$\mathbf{Y}_{\mathrm{ref}} = [\mathbf{Y}_{\mathrm{mid}}; \mathbf{Y}_{\rho}; \mathbf{Y}_{\theta}] = \mathcal{B}_{\mathrm{ref}}(\mathbf{F}) \in \mathbb{R}^{|\mathrm{D}_N|\times H\times W\times 3}, \tag{6}$$

where $\mathbf{Y}_{\mathrm{mid}}$ provides a probability for each spatial location being the midpoint of a reflection axis, and $\mathbf{Y}_{\rho}$ and $\mathbf{Y}_{\theta}$ are the regression outputs for the axis length and orientation at that location. To obtain reflection component map, we apply pooling across the group dimension, Pool_G as follows:

$$\mathbf{O}_{\mathrm{ref}} = \mathrm{Pool}_G(\mathbf{Y}_{\mathrm{ref}}) \in \mathbb{R}^{H\times W\times 3}. \tag{7}$$

At each position (x, y), a reflection axis prediction is parameterized as (x, y, p, ρ, θ), where p is the midpoint probability, ρ the length, and θ the orientation. The start and end points of the predicted axis are given by:

$$\begin{bmatrix} x_{\mathrm{s}} \\ y_{\mathrm{s}} \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \frac{\rho}{2} \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}, \tag{8}$$

$$\begin{bmatrix} x_{\mathrm{e}} \\ y_{\mathrm{e}} \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} - \frac{\rho}{2} \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix}. \tag{9}$$

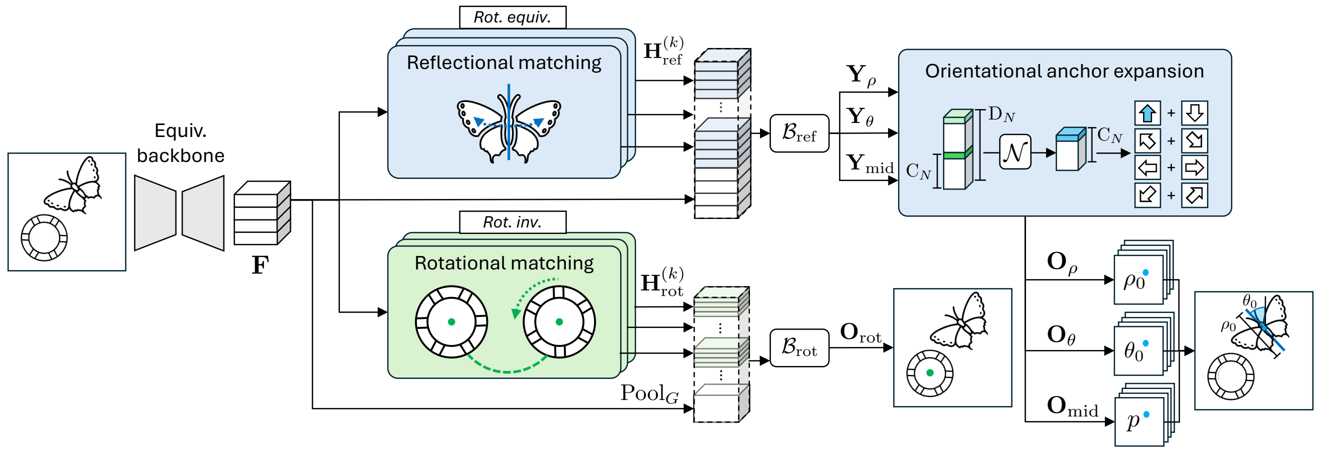

Figure 2. **Overall architecture of our proposed instance-level symmetry detection network.** Given an input image, a D_N-equivariant backbone extracts features $\mathbf{F} \in \mathbb{R}^{H\times W\times C|\mathrm{D}_N|}$. The reflection branch (top) employs equivariant reflectional matching and orientational anchor expansion to predict reflection axes as parameterized line segments $(\alpha, x, y, p, \rho, \theta)$. The rotation branch (bottom) applies invariant rotational matching to detect rotation axes and classify their fold classes parameterized as (x, y, p_c).

Rotation symmetry detection. For rotation symmetry, our goal is to predict the positions of rotation centers and classify their fold (symmetry order). An n-fold rotational symmetry means the pattern looks the same after rotation by $\frac{2\pi}{n}$ (for example, a 4-fold symmetry repeats every $\frac{\pi}{2}$). To predict both axis existance and fold class, rotation branch $\mathcal{B}_{\text{rot}}$ produces the multi-class classification score map:

$$\mathbf{O}_{\text{rot}} = \mathcal{B}_{\text{rot}}(\text{Pool}_G(\mathbf{F})) \in \mathbb{R}^{H\times W\times S}, \tag{10}$$

where S is the number of fold classes including the background class. Each rotation axis prediction is represented as (x, y, p_s), with p_s as the probability of the s-th fold class.

Training objective. The training objective includes both reflection and rotation symmetry losses. For reflection symmetry, we apply losses for midpoint classification, length regression, and orientation regression. Midpoint classification is optimized using weighted binary cross-entropy:

$$\mathcal{L}_{\text{mid}} = \mathbb{E}_{(x,y)}[-\gamma_{\text{ref}}\, p\log(\hat{p}) - (1-p)\log(1-\hat{p})], \tag{11}$$

where p is the ground truth label, $\hat{p}$ is the predicted probability, and γ_{ref} is a weighting factor. Length ρ and orientation θ regression losses are applied only at positions with valid midpoints ($p = 1$), enforced by the indicator function $\mathbb{I}_{p=1}$:

$$\mathcal{L}_\rho = \mathbb{E}_{(x,y)}[\mathbb{I}_{p=1} \cdot \text{SmoothL1}(\rho, \hat{\rho})], \tag{12}$$

$$\mathcal{L}_\theta = \mathbb{E}_{(x,y)}[\mathbb{I}_{p=1} \cdot |\theta - \hat{\theta}|], \tag{13}$$

where ρ and θ denote ground truth values, and $\hat{\rho}$ and $\hat{\theta}$ are the predicted values. For rotation symmetry, fold classification is optimized using weighted multi-class cross-entropy:

$$\mathcal{L}_{\text{fold}} = \mathbb{E}_{(x,y,s)}[-\gamma_{\text{rot}}\, p_s \log \hat{p}_s], \tag{14}$$

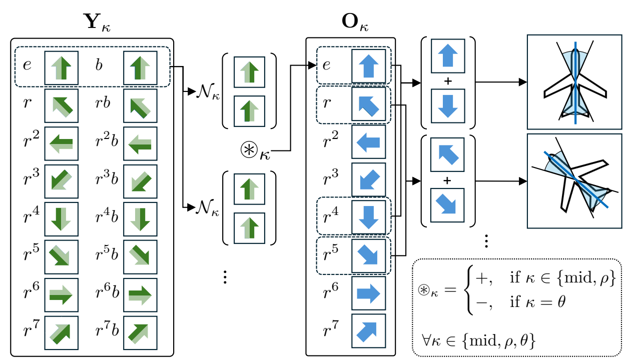

Figure 3. **Illustration of our orientational anchor expansion on $\mathbf{D}_8$ group.** The D_8-equivariant features $\mathbf{Y}_\kappa$ undergo transformation $\mathcal{N}_\kappa$ and aggregation $\circledast_\kappa$, creating C_8-equivariant features $\mathbf{O}_\kappa$. These are combined across opposite orientations to handle the θ and $\theta + \pi$ equivalence, allowing each orientation channel to specialize in specific angular ranges and improve detection of axes with overlapping midpoints. Each arrow represents a feature map.

where s represents the fold class and γ_{rot} is applied at positions with ground truth rotation axis and correct fold class ($p = 1$). The total loss is a weighted sum of loss terms:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{mid}} + \lambda_\rho\mathcal{L}_\rho + \lambda_\theta\mathcal{L}_\theta + \lambda_{\text{fold}}\mathcal{L}_{\text{fold}}. \tag{15}$$

4.2. Orientational anchor expansion

In standard object detection, anchor boxes [33, 41] are placed at various scales and aspect ratios to guide bounding box regression. Analogously, our reflection symmetry branch initially treats each pixel as an anchor for a potential symmetry axis and learns to regress the orientation and length of the axis. However, this straightforward approach does not fully exploit the orientation dimension provided by our group-equivariant features. To address this, we introduce orientational anchor expansion, integrating the group

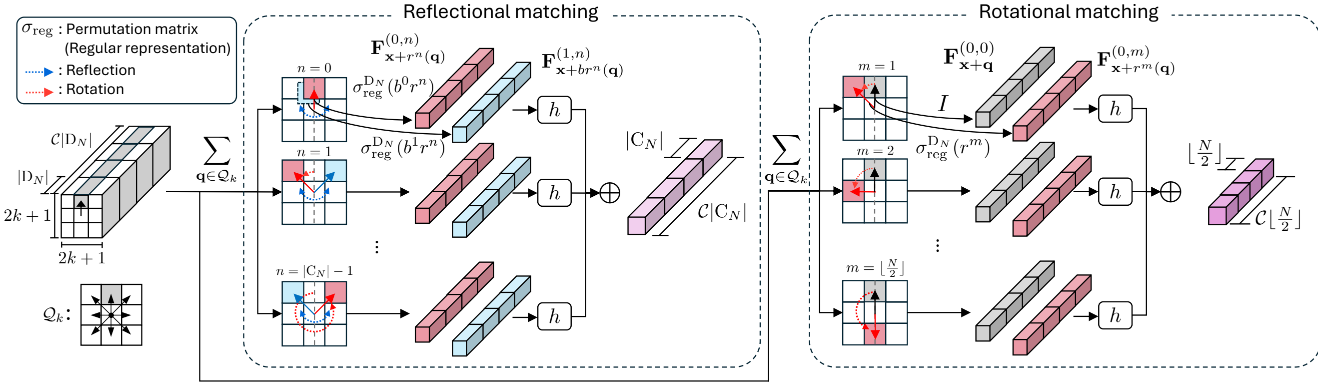

Figure 4. **Illustration of our equivariant reflectional matching (left) and invariant rotational matching (right) modules.** The reflectional matching computes similarity scores between rotated features and their reflections across all $|\mathrm{C}_N|$ rotation angles, preserving dihedral group equivariance with rotation invariance. The rotational matching computes similarities between feature pairs with different rotation angle interval, yielding rotation-invariant features for detecting n-fold rotation symmetry centers. Both modules incorporate spatial neighborhoods $\mathcal{Q}_k$ for robust detection across multiple scales.

dimension into the detection framework for orientationally specialized axis detection and improved handling of axes with overlapping midpoints but different orientations. Fig. 3 illustrates the proposed orientational anchor expansion.

Reflectional counterpart aggregation. In a D_N-equivariant feature map, each group dimensional channel of $\mathbf{F}$ corresponds to a particular element of the dihedral group. Recall that reflection branch outputs in Eq. (6) produce tensors $\mathbf{Y}_\kappa \in \mathbb{R}^{|\mathrm{D}_N| \times H \times W}$ for each component $\kappa \in \{\text{mid}, \rho, \theta\}$. The $2N$ channels in this first dimension can be thought of as N pairs, where each pair $(i,\ i+N)$ consists of a feature responding to some rotation r^i and its reflected version br^i. To make use of the orientation-specific information, we aggregate each such pair of reflection counterparts into a single response, in a way that preserves the feature's equivariance under pure rotations.

For the midpoint score $\mathbf{Y}_{\text{mid}}$ and length $\mathbf{Y}_\rho$, which are unchanged by reflecting the image, we add the two responses. For the orientation output $\mathbf{Y}_\theta$, which flips sign under reflection (an axis at angle θ becomes $-\theta$), we subtract the reflected response. Formally, let $\mathbf{Y}_\kappa^{(i)}$ denote the feature map for the i-th rotation channel and $\mathbf{Y}_\kappa^{(i+N)}$ the corresponding reflection channel. We compute an aggregated feature $\tilde{\mathbf{Y}}_\kappa$ with only N channels as:

$$\tilde{\mathbf{Y}}_\kappa = \bigoplus_{i=1}^{|\mathrm{C}_N|} \left[\mathcal{N}_\kappa([\mathbf{Y}_\kappa^{(i)}; \mathbf{Y}_\kappa^{(i+N)}]) \circledast_\kappa \mathcal{N}_\kappa([\mathbf{Y}_\kappa^{(i+N)}; \mathbf{Y}_\kappa^{(i)}]) \right], \quad (16)$$

where $\circledast_\kappa$ is the pairwise aggregation operator defined as $\circledast_\kappa = +$ for $\kappa \in \{\text{mid}, \rho\}$ and $\circledast_\kappa = -$ for $\kappa = \theta$. Rather than directly adding or subtracting feature maps which discards useful details, we apply a learnable transformation $\mathcal{N}_\kappa$ to extract and reweight information from each channel before combining, while preserving both the reflection transformation properties and C_N-equivariance.

Orientational anchor construction. Even after merging reflection pairs, there remains an ambiguity in the orientation representation: a line at orientation θ is equivalent to the one at $\theta + \pi$, since both describe the same physical axis line. To address this ambiguity, we combine the aggregated response at rotation channel index α with that at $\alpha + {}^{N}\!/_{2}$ to produce $\mathbf{O}_{\text{ref}} = [\mathbf{O}_{\text{mid}}; \mathbf{O}_\rho; \mathbf{O}_\theta] \in \mathbb{R}^{|\mathrm{C}_N|/2 \times H \times W \times 3}$:

$$\mathbf{O}_{\kappa,\alpha} = \tilde{\mathbf{Y}}_{\kappa,\alpha} + \tilde{\mathbf{Y}}_{\kappa,\alpha+N/2}, \quad \alpha = 1, \ldots, \tfrac{N}{2}, \quad (17)$$

for each component $\kappa \in \{\text{mid}, \rho, \theta\}$. Each anchor $\mathbf{O}_\alpha$ specializes in detecting axes with orientation offsets within $[-\frac{\pi}{N}, \frac{\pi}{N})$ from its base orientation $\frac{2\pi\alpha}{N}$. We predict offsets rather than absolute orientations to directly adapt invariant orientation regression values across different anchor orientations. At each position (α, x, y), an axis is represented as $(\alpha, x, y, p, \rho, \theta)$, where the output $\mathbf{O}_{(\alpha,x,y)} = (p, \rho, \theta)$ determines its start and end points as:

$$\begin{bmatrix} x_{\text{s},\alpha} \\ y_{\text{s},\alpha} \end{bmatrix} = \begin{bmatrix} x_\alpha \\ y_\alpha \end{bmatrix} + \frac{\rho}{2} \begin{bmatrix} \cos(\theta_\alpha) \\ \sin(\theta_\alpha) \end{bmatrix}, \quad (18)$$

$$\begin{bmatrix} x_{\text{e},\alpha} \\ y_{\text{e},\alpha} \end{bmatrix} = \begin{bmatrix} x_\alpha \\ y_\alpha \end{bmatrix} - \frac{\rho}{2} \begin{bmatrix} \cos(\theta_\alpha) \\ \sin(\theta_\alpha) \end{bmatrix}, \quad (19)$$

where $\theta_\alpha = \frac{2\pi\alpha}{N} + \theta$ represents the absolute orientation.

4.3. Reflectional matching

Reflection symmetry can be validated by comparing a pattern with its mirrored counterpart, known as reflectional matching [3, 37]. Unlike hand-crafted descriptors such as SIFT [36], conventional neural features [13, 20] lack rotation and reflection equivariance, limiting their effectiveness. To address this, we leverage D_N-equivariant features [5] for reflectional matching, providing a strong cue for symmetry detection. Fig. 4 (left) illustrates the detailed process.

For a feature fiber $\mathbf{f} \in \mathbb{R}^{\mathcal{C}|\mathrm{D}_N|}$ from a D_N-equivariant feature map $\mathbf{F}$, its transformation under l reflections and n

rotations is:

$$\mathbf{f}^{(l,n)} = \bigoplus_{c=1}^{\mathcal{C}} \sigma_{\text{reg}}^{\text{D}_N}(b^l r^n)\mathbf{f}_c \in \mathbb{R}^{\mathcal{C}|\text{D}_N|}, \quad (20)$$

where $\mathbf{f}_c \in \mathbb{R}^{|\text{D}_N|}$ represents the group-equivariant subset of the fiber, and $\mathbf{f} = [\mathbf{f}_1^\top, \ldots, \mathbf{f}_{\mathcal{C}}^\top]^\top$. Here, $\sigma_{\text{reg}}^{\text{D}_N}(b^l r^n)$ denotes the regular representation of D_N for l reflections and n rotations. The group-aware similarity h between two fibers $\mathbf{f}^1, \mathbf{f}^2 \in \mathbb{R}^{\mathcal{C}|\text{D}_N|}$ is defined as:

$$h(\mathbf{f}^1, \mathbf{f}^2) = \bigoplus_{c=1}^{\mathcal{C}} \frac{\mathbf{f}_c^1 \cdot \mathbf{f}_c^2}{\|\mathbf{f}_c^1\|\|\mathbf{f}_c^2\|} \in \mathbb{R}^{\mathcal{C}}. \quad (21)$$

To capture symmetry across orientations, reflectional similarity scores are computed for each rotation, comparing rotated and rotated-then-reflected fibers:

$$\mathbf{H}_{\text{ref},\mathbf{x}} = \bigoplus_{n=0}^{|\text{C}_N|-1} h(\mathbf{F}_{\mathbf{x}}^{(0,n)}, \mathbf{F}_{\mathbf{x}}^{(1,n)}) \in \mathbb{R}^{\mathcal{C}|\text{C}_N|}, \quad (22)$$

where $\mathbf{F}_{\mathbf{x}}^{(0,n)}$ and $\mathbf{F}_{\mathbf{x}}^{(1,n)}$ represent fibers at position $\mathbf{x}$ under the regular representation for n rotations, with and without reflection. The resulting similarity score map $\mathbf{H} \in \mathbb{R}^{\mathcal{C}|\text{C}_N| \times H \times W}$ is equivariant under the dihedral group while remaining reflection-invariant. To detect broader symmetries beyond single points, we extend matching to spatial neighborhoods defined by a set of 2D offset vectors:

$$\mathcal{Q}_k = \{(i,j) \mid i,j \in \{-k, \ldots, k\}\}, \quad (23)$$

where $k \in \mathbb{N}$ controls the neighborhood size. The neighborhood similarity is computed as:

$$\mathbf{H}_{\text{ref},\mathbf{x}}^{(k)} = \sum_{\mathbf{q} \in \mathcal{Q}_k} \bigoplus_{n=0}^{|\text{C}_N|-1} h(\mathbf{F}_{\mathbf{x}+r^n(\mathbf{q})}^{(0,n)}, \mathbf{F}_{\mathbf{x}+br^n(\mathbf{q})}^{(1,n)}) \in \mathbb{R}^{\mathcal{C}|\text{C}_N|}, \quad (24)$$

where $b^l r^n(\mathbf{q})$ denotes the transformed offset after n rotations and l reflections. To improve robustness, we use multi-scale reflectional similarity features $\mathbf{H}_{\text{ref}}^{(k_1)}, \ldots, \mathbf{H}_{\text{ref}}^{(k_M)}$, concatenated with the base feature map $\mathbf{F}$ to capture symmetry across various spatial scales while preserving equivariance. The matching output is equivariant to D_N while preserving reflection invariance, as demonstrated in the detailed proof provided in the Supplementary Material.

4.4. Rotational matching

Rotation symmetry is identified by comparing a pattern with its rotated version around its axis. Our rotational matching module implements this by comparing features with their rotated versions around each candidate center point (Fig. 4 right). An n-fold rotational symmetry remains invariant under every $\frac{2\pi}{n}$ rotation. To reduce redundancy in similarity comparisons, we exploit the consistency of feature comparisons at fixed angular separations, requiring only $\lfloor \frac{N}{2} \rfloor$ unique comparisons instead of ${}_N\text{C}_2$ feature pairs. The complete rotational matching feature is computed as:

$$\mathbf{H}_{\text{rot},\mathbf{x}} = \bigoplus_{m=1}^{\lfloor \frac{N}{2} \rfloor} h(\mathbf{F}_{\mathbf{x}}^{(0,0)}, \mathbf{F}_{\mathbf{x}}^{(0,m)}) \in \mathbb{R}^{\mathcal{C}\lfloor \frac{N}{2} \rfloor}, \quad (25)$$

which remains dihedral group-invariant, as similarity values are preserved. To extend matching to spatial neighborhoods, we use the approach from Sec. 4.3:

$$\mathbf{H}_{\text{rot},\mathbf{x}}^{(k)} = \sum_{\mathbf{q} \in \mathcal{Q}_k} \bigoplus_{m=1}^{\lfloor \frac{N}{2} \rfloor} h(\mathbf{F}_{\mathbf{x}+\mathbf{q}}^{(0,0)}, \mathbf{F}_{\mathbf{x}+r^m(\mathbf{q})}^{(0,m)}) \in \mathbb{R}^{\mathcal{C}\lfloor \frac{N}{2} \rfloor}. \quad (26)$$

Following the multi-scale approach in Sec. 4.3, we compute features at multiple kernel sizes and concatenate them with the pooled base feature map from Sec. 4.1, enabling precise detection of rotation axis and fold classes. The resulting output remains invariant to both rotation and reflection.

5. Experiments

5.1. Implementation details

Dataset. We use DENDI dataset [44], which provides annotations for reflection symmetry axes and rotation symmetry centers with folds. Data augmentation includes flipping, rotation, and color jittering. We extract 7k axis-annotated and 8k rotation center-annotated masks from the training set, pasting them onto other images without overlapping existing annotations. Adding 1-6 objects per iteration and repeating three times expands the dataset to approximately 30k training images. We also evaluate F1-scores on reflection symmetry datasets SDRW [32] and LDRS [44] for comparison with previous methods.

Evaluation metrics. For reflection symmetry, structural Average Precision (sAP) [62] is adopted, where a predicted axis is considered a true positive if $d_1^2 + d_2^2 < \tau$ or $d_{\text{mid}}^2 < \frac{\tau}{2}$ with at least 70% overlap within an annotated ellipse. Here, d_1 and d_2 are endpoint distances between ground truth and predicted axis, and d_{mid} is the distance between ground truth ellipse center and predicted axis midpoint. For rotation symmetry, we evaluate sAP for the rotation axis with threshold $d_{\text{center}}^2 < \frac{\tau}{2}$ and fold sAP, which also requires correct fold class prediction. Here, d_{center} denotes the distance between predicted and ground truth rotation axes. Results are reported at $\tau = 5, 10, 15$ pixels for both tasks. For comparison with heatmap-based methods, we report F1-scores after generating heatmaps by dilating ground-truth and predicted symmetry axes by 5 pixels. [15, 44].

ICCV #16

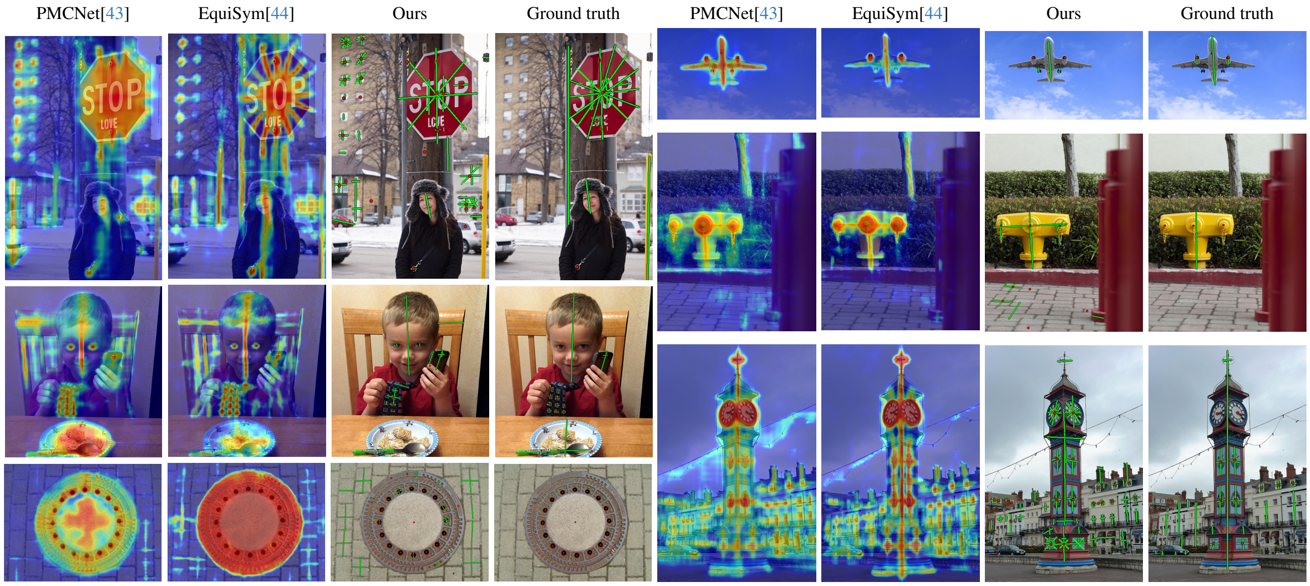

Figure 5. **Qualitative comparison of symmetry detection methods.** Our instance-wise approach produces clearer, more precise symmetry instances compared to heatmap-based methods [43, 44], especially for smaller objects and complex scenes. Green lines in ground truth and our results represent reflection axes, while red points represent rotation axes.

Model and training. We use a D_8-equivariant ResNet-34 [5, 20] as the feature extractor. Both reflectional and rotational matching modules employ multi-scale similarity feature (1, 3, and 5). In the reflection branch, group convolution [5] is implemented by rotating the image, permuting group dimension channels, and applying standard convolution. We chose this approach because the `e2cnn` [54] framework does not natively support group convolution for reflection-invariant dihedral groups, nor various operations such as deformable convolution [11] while preserving equivariance. The model is trained for 100 epochs with a batch size of 32 using AdamW [26], starting with a learning rate of 10^{-3}, which decreases $\frac{1}{10}$ at epochs 50 and 75. Loss weights are set as $\lambda_\rho = 1$ for length, $\lambda_\theta = 150$ for orientation (to account for the radian scale), and $\lambda_{\text{fold}} = 2$ for rotation fold classification. Weighted binary cross-entropy with a positive class weight of 3 is applied to $\mathcal{L}_{\text{mid}}$ and $\mathcal{L}_{\text{fold}}$.

5.2. Evaluation of the proposed method

Reflection symmetry detection. As shown in the last row of the Tab. 1, the proposed model achieves sAP scores of 18.7%, 22.7%, and 24.7% at 5, 10, and 15-pixel thresholds on the DENDI dataset [44]. Fig. 5 demonstrates robust reflection symmetry detection across diverse scenes, handling multiple orientations and scales, even in complex backgrounds. The orientational anchor expansion resolves overlapping midpoints, addressing a key challenge in axis-level detection.

Rotation symmetry detection. The rotation symmetry branch outputs classification scores for multiple folds. For center detection, we pool the maximum score as the center probability for binary evaluation. In the last row of the Tab. 2, we report both center sAP and fold sAP. Our method achieves center sAP scores of 36.8%, 39.1%, and 40.0% and fold AP scores of 26.6%, 28.3%, and 28.9% at 5, 10, and 15-pixel thresholds, respectively. Fold misclassifications primarily occur between 2-fold and 4-fold symmetries (*e.g.*, rectangles vs. squares). Fig. 5 shows robust detection across complex scenes.

Method	Ref. sAP (%) @5	@10	@15
Axis-level detection	6.2	9.3	11.2
+ Orientational anchors	16.6	19.9	21.1
+ Ref. match$_{k=0}$	17.6	20.7	21.8
+ Ref. match$_{k=0,1}$	18.4	22.0	23.7
+ Ref. match$_{k=0,1,2}$	**18.8**	**22.7**	**24.7**

Table 1. Ablation results for reflection symmetry detection on the DENDI dataset. Orientational anchors denotes our anchor expansion approach, and Ref. match$_k$ represents reflectional matching with kernel sizes $2k+1$. Best results are shown in **bold**.

5.3. Ablation study

Reflection symmetry detection. We conduct ablation studies on our reflection symmetry components in Tab. 1. The baseline axis-level detection achieves sAP scores of 6.2%, 9.3%, and 11.2%. Adding orientational anchors significantly improves performance, increasing sAP to 16.6%, 19.9%, and 21.1% by enabling orientation-specific feature learning. Incorporating single-kernel reflectional matching

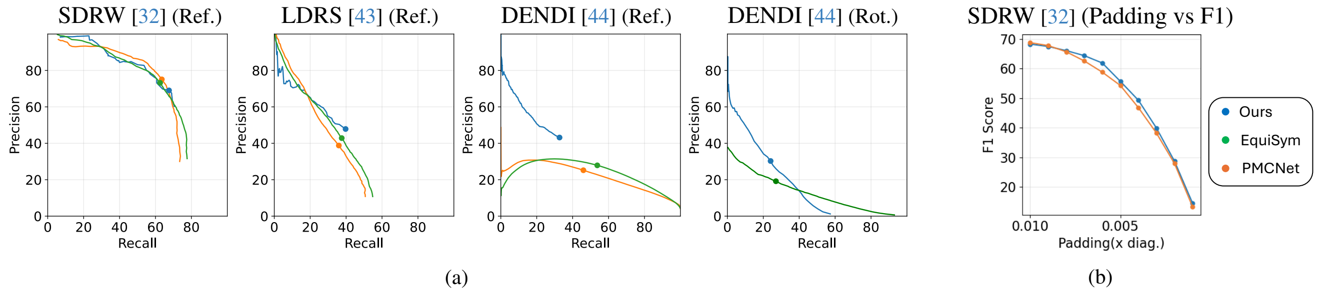

Figure 6. Evaluation of our symmetry detection approach: (a) Precision-recall curves for reflection symmetry detection on SDRW [32], LDRS [43], and DENDI [44] datasets, and rotation symmetry detection on the DENDI dataset; (b) Analysis of F1-score with varying axis padding values(multiplied to image diagonal) on the SDRW dataset. For rotation symmetry evaluation on DENDI, we compare only with EquiSym [44], as PMCNet [43] does not support rotation symmetry detection.

Method	Center sAP (Fold sAP) (%)		
	@5	@10	@15
Axis-level detection	31.5(22.5)	34.7(24.6)	35.7(25.3)
+ Rot. match$_{k=0}$	35.9(25.4)	37.8(26.6)	37.0(27.2)
+ Rot. match$_{k=0,1}$	36.2(26.2)	38.2(27.8)	37.4(28.1)
+ Rot. match$_{k=0,1,2}$	**36.8(26.6)**	**39.1(28.3)**	**40.0(28.9)**

Table 2. Ablation results for rotation symmetry detection on the DENDI dataset. Rot. match$_k$ represents rotational matching with kernel sizes $2k+1$. Best results are shown in **bold**.

($k = 0$) further boosts performance, achieving sAP scores of 17.6%, 20.7%, and 21.8%. Expanding to multi-kernel matching ($k = 0, 1$) enhances detection, reaching 18.4%, 22.0%, and 23.7%. The best performance is obtained with $k = 0, 1, 2$, achieving sAP scores of 18.8%, 22.7%, and 24.7%. These results confirm that multi-scale reflectional matching effectively captures symmetry patterns across different spatial scales.

Rotation symmetry detection. Tab. 2 shows the effectiveness of our rotational matching approach. The baseline axis-level detection achieves sAP scores of 31.5%, 34.7%, and 35.7%, with fold sAP scores of 22.5%, 24.6%, and 25.3%. Adding single-kernel rotational matching ($k = 0$) improves sAP by 4.4%, 3.1%, and 1.3%, with greater gains at smaller thresholds (5 pixels), indicating better localization. Expanding to multi-kernel matching ($k = 0, 1$ and $k = 0, 1, 2$) further enhances performance. Our final model achieves sAP scores of 36.8%, 39.1%, and 40.0%, with fold sAP scores of 26.6%, 28.3%, and 28.9%, demonstrating the effectiveness of rotational matching across scales.

5.4. Comparison with the state-of-the-art methods

F1-score. Tab. 3 shows F1-scores across multiple benchmarks. Our method outperforms previous work on LDRS [43] (+3.4%) and DENDI [44] (+0.5%) for reflection symmetry and achieves a significant gain(+4.4%) in rotation symmetry detection on DENDI. For the SDRW [32], our method (68.3%) is comparable to PMCNet (68.8%). This

Method	Ref. F1 (%)			Rot. F1 (%)
	SDRW [32]	LDRS [43]	DENDI [44]	DENDI [44]
PMCNet [43]	**68.8**	37.3	32.6	-
EquiSym [44]	67.5	40.0	36.7	22.4
Ours	68.3	**43.4**	**37.2**	**26.8**

Table 3. Comparison with the state-of-the-art methods using pixel-wise F1-score on multiple datasets. Ref. and Rot. denote reflection and rotation symmetry respectively. Best results in **bold**.

slight difference stems from the disparity between heatmap-based segmentation and detection approaches. As shown in Fig. 6(b), when the evaluation criterion becomes more stringent (smaller padding values), our axis-level approach outperforms PMCNet's region-based predictions due to more precise axis localization.

PR curve. Precision-recall curves in Fig. 6(a) further highlight performance differences. Our method maintains higher precision, especially in LDRS and DENDI. Unlike pixel-level methods that boost recall by predicting all pixels, our approach models symmetry as geometric primitives, where a single midpoint score affects the entire axis. This results in higher precision but more variable recall. Post-processing steps like Non-Maximum Suppression and score thresholding further prioritize precision over recall.

6. Conclusion

We have introduced a dihedral group-equivariant approach for axis-level symmetry detection, representing symmetries as geometric primitives instead of pixel-level heatmaps. Our method integrates orientational anchors and reflectional matching for reflection symmetry detection, and invariant rotational matching for rotation symmetry detection to capture symmetry across orientations and scales. Experiments demonstrate superior performance over existing methods, with ablations validating the effectiveness of our approach. Future work can extend our model to continuous groups, 3D spaces, and varying viewpoints for real-world applications.

References

[1] Habib Akbar, Muhammad Munwar Iqbal, Abid Ali, Amna Parveen, Nagwan Abdel Samee, Manal Abdullah Alohali, and Mohammed Saleh Ali Muthanna. Detecting rotational symmetry in polar domain based on sift. *IEEE Access*, 11: 68643–68652, 2023. 2

[2] Georg Bökman, Johan Edstedt, Michael Felsberg, and Fredrik Kahl. Steerers: A framework for rotation equivariant keypoint descriptors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4885–4895, 2024. 2

[3] Minsu Cho and Kyoung Mu Lee. Bilateral symmetry detection via symmetry-growing. In *BMVC*, pages 1–11. Citeseer, 2009. 1, 2, 5

[4] Marcelo Cicconet, David GC Hildebrand, and Hunter Elliott. Finding mirror symmetry via registration and optimal symmetric pairwise assignment of curves. pages 1749–1758, 2017. 2

[5] Taco Cohen and Max Welling. Group equivariant convolutional networks. In *International conference on machine learning*, pages 2990–2999. PMLR, 2016. 2, 3, 5, 7

[6] Taco Cohen, Mario Geiger, and Maurice Weiler. A general theory of equivariant cnns on homogeneous spaces. *arXiv preprint arXiv:1811.02017*, 2018. 2

[7] Taco Cohen, Maurice Weiler, Berkay Kicanaoglu, and Max Welling. Gauge equivariant convolutional networks and the icosahedral cnn. In *International Conference on Machine Learning*, pages 1321–1330. PMLR, 2019. 2

[8] Taco S Cohen and Max Welling. Steerable cnns. *arXiv preprint arXiv:1612.08498*, 2016. 2, 3

[9] Taco S Cohen, Mario Geiger, and Maurice Weiler. Intertwiners between induced representations (with applications to the theory of equivariant neural networks). *arXiv preprint arXiv:1803.10743*, 2018. 2

[10] Hugo Cornelius, Michal Perďoch, Jiří Matas, and Gareth Loy. Efficient symmetry detection using local affine frames. In *Proceedings of the 15th Scandinavian Conference on Image Analysis (SCIA)*, pages 152–161. Springer, 2007. 2

[11] Jifeng Dai, Haozhi Qi, Yuwen Xiong, Yi Li, Guodong Zhang, Han Hu, and Yichen Wei. Deformable convolutional networks. In *Proceedings of the IEEE international conference on computer vision*, pages 764–773, 2017. 7

[12] Sander Dieleman, Kyle W Willett, and Joni Dambre. Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Monthly notices of the royal astronomical society*, 450(2):1441–1459, 2015. 2

[13] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020. 5

[14] Kunihiko Fukushima and Masayuki Kikuchi. Symmetry axis extraction by a neural network. *Neurocomputing*, 69(16-18): 1827–1836, 2006. 1, 2

[15] Christopher Funk and Yanxi Liu. Beyond planar symmetry: Modeling human perception of reflection and rotation symmetries in the wild. In *ICCV*, pages 793–803, 2017. 1, 2, 3, 6

[16] Robert Gens and Pedro M Domingos. Deep symmetry networks. In *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2014. 2

[17] Martin Giurfa, Bettina Eichmann, and Randolf Menzel. Symmetry perception in an insect. *Nature*, 382(6590):458–461, 1996. 1

[18] Alessandro Gnutti, Fabrizio Guerrini, and Riccardo Leonardi. Combining appearance and gradient information for image symmetry detection. *IEEE TIP*, 2021. 2

[19] Jiaming Han, Jian Ding, Nan Xue, and Gui-Song Xia. Redet: A rotation-equivariant detector for aerial object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2786–2795, 2021. 2

[20] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016. 3, 5, 7

[21] Emiel Hoogeboom, Jorn WT Peters, Taco S Cohen, and Max Welling. Hexaconv. *arXiv preprint arXiv:1803.02108*, 2018. 2

[22] Kun Huang, Yifan Wang, Zihan Zhou, Tianjiao Ding, Shenghua Gao, and Yi Ma. Learning to parse wireframes in images of man-made environments. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 626–635, 2018. 3

[23] Cornelius Hugo and Loy Gareth. Detecting bilateral symmetry in perspective. In *CVPRW*, 2006. 2

[24] Jihyeon Je, Jiayi Liu, Guandao Yang, Boyang Deng, Shengqu Cai, Gordon Wetzstein, Or Litany, and Leonidas Guibas. Robust symmetry detection via riemannian langevin dynamics. In *SIGGRAPH Asia 2024 Conference Papers*, pages 1–11, 2024. 2, 3

[25] Yosi Keller and Yoel Shkolnisky. A signal processing approach to symmetry detection. *IEEE TIP*, 15(8):2198–2207, 2006. 1, 2

[26] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *ICLR*, 2015. 7

[27] Nahum Kiryati and Yossi Gofman. Detecting symmetry in grey level images: The global optimization approach. *International Journal of Computer Vision*, 29(1):29–45, 1998. 1

[28] Jongmin Lee and Minsu Cho. 3d equivariant pose regression via direct wigner-d harmonics prediction. *Advances in Neural Information Processing Systems*, 37:127765–127798, 2024. 2

[29] Seungkyu Lee and Yanxi Liu. Skewed rotation symmetry group detection. *IEEE TPAMI*, 32(9):1659–1672, 2009. 2

[30] Seungkyu Lee, Robert T Collins, and Yanxi Liu. Rotation symmetry group detection via frequency analysis of frieze-expansions. In *CVPR*, pages 1–8. IEEE, 2008. 1, 2

[31] Xiang Li, Zixuan Huang, Anh Thai, and James M Rehg. Symmetry strikes back: From single-image symmetry detection to 3d generation. *arXiv preprint arXiv:2411.17763*, 2024. 2, 3

[32] Jingchen Liu, George 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 *CVPRW*, pages 200–205, 2013. 6, 8

[33] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. Ssd: Single shot multibox detector. In *European Conference on Computer Vision (ECCV)*, pages 21–37. Springer, 2016. 4

[34] Yanxi Liu, Robert T Collins, and Yanghai Tsin. A computational model for periodic pattern perception based on frieze and wallpaper groups. *IEEE TPAMI*, 26(3):354–371, 2004. 2

[35] Yanxi Liu, Hagit Hel-Or, and Craig S Kaplan. *Computational symmetry in computer vision and computer graphics*. Now publishers Inc, 2010. 1

[36] David G Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004. 1, 2, 5

[37] Gareth Loy and Jan-Olof Eklundh. Detecting symmetry and symmetric constellations of features. In *ECCV*, pages 508–521. Springer, 2006. 1, 2, 5

[38] Diego Marcos, Michele Volpi, Nikos Komodakis, and Devis Tuia. Rotation equivariant vector field networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 5048–5057, 2017. 2

[39] Anders Pape Møller and Randy Thornhill. Bilateral symmetry and sexual selection: a meta-analysis. *The American Naturalist*, 151(2):174–192, 1998. 1

[40] V Shiv Naga Prasad and Larry S Davis. Detecting rotational symmetries. In *ICCV*, pages 954–961. IEEE, 2005. 1, 2

[41] Shaoqing Ren. Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*, 2015. 4

[42] J.J. Rotman. An introduction to the theory of groups. *Graduate Texts in Mathematics*, 2012. 2

[43] Ahyun Seo, Woohyeon Shim, and Minsu Cho. Learning to discover reflection symmetry via polar matching convolution. In *ICCV*, 2021. 1, 2, 3, 7, 8

[44] Ahyun Seo, Byungjin Kim, Suha Kwak, and Minsu Cho. Reflection and rotation symmetry detection via equivariant learning. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 1, 2, 3, 6, 7, 8

[45] Dinggang Shen, Horace HS Ip, and Eam Khwang Teoh. Robust detection of skewed symmetries by combining local and semi-local affine invariants. *PR*, 34(7):1417–1428, 2001. 2

[46] Sinha Sudipta N., Ramnath Krishnan, and Szeliski Richard. Detecting and reconstructing 3d mirror symmetric objects. In *ECCV*, pages 586–600. Springer, 2012. 2

[47] Changming Sun. Symmetry detection using gradient information. *Pattern Recognition Letters*, 16(9):987–996, 1995. 2

[48] Ching L Teo, Cornelia Fermuller, and Yiannis Aloimonos. Detection and segmentation of 2d curved reflection symmetric structures. In *ICCV*, pages 1644–1652, 2015. 1, 2

[49] Matthieu Terris, Thomas Moreau, Nelly Pustelnik, and Julian Tachella. Equivariant plug-and-play image reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 25255–25264, 2024. 2

[50] Stavros Tsogkas and Iasonas Kokkinos. Learning-based symmetry detection in natural images. In *ECCV*, pages 41–54. Springer, 2012. 1, 2

[51] Christopher W Tyler. Empirical aspects of symmetry perception. *Spatial Vision*, 9(1):1–7, 1995. 1

[52] Johan Wagemans. Detection of visual symmetries. *Spatial vision*, 9(1):9–32, 1995. 1

[53] Zhaozhong Wang, Lianrui Fu, and YF Li. Unified detection of skewed rotation, reflection and translation symmetries from affine invariant contour features. *PR*, 47(4):1764–1776, 2014. 2

[54] Maurice Weiler and Gabriele Cesa. General E(2)-Equivariant Steerable CNNs. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2019. 2, 7

[55] Maurice Weiler, Mario Geiger, Max Welling, Wouter Boomsma, and Taco Cohen. 3d steerable cnns: Learning rotationally equivariant features in volumetric data. *arXiv preprint arXiv:1807.02547*, 2018. 2, 3

[56] Max Wertheimer. Laws of organization in perceptual forms. In *A source book of Gestalt psychology*, pages 71–88. Routledge & Kegan Paul, London, 1938. 1

[57] Daniel E Worrall, Stephan J Garbin, Daniyar Turmukhambetov, and Gabriel J Brostow. Harmonic networks: Deep translation and rotation equivariance. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5028–5037, 2017. 2

[58] Yinshuang Xu, Dian Chen, Katherine Liu, Sergey Zakharov, Rareș Ambruș, Kostas Daniilidis, and Vitor Guizilini. $se(3)$ equivariant ray embeddings for implicit multi-view depth estimation. *Advances in Neural Information Processing Systems*, 37:13627–13659, 2025. 2

[59] Nan Xue, Tianfu Wu, Song Bai, Fudong Wang, Gui-Song Xia, Liangpei Zhang, and Philip HS Torr. Holistically-attracted wireframe parsing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2788–2797, 2020. 3

[60] Haotian Zhang, Yicheng Luo, Fangbo Qin, Yijia He, and Xiao Liu. Elsd: Efficient line segment detector and descriptor. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2969–2978, 2021. 1, 3

[61] Zhaoxuan Zhang, Bo Dong, Tong Li, Felix Heide, Pieter Peers, Baocai Yin, and Xin Yang. Single depth-image 3d reflection symmetry and shape prediction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 8896–8906, 2023. 2

[62] Yichao Zhou, Haozhi Qi, and Yi Ma. End-to-end wireframe parsing. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 962–971, 2019. 3, 6