

# PHASE-PRESERVING ANALYTICAL FEATURES FROM SOLID HARMONIC WAVELET BISPECTRUM SIMPLIFY DECISION BOUNDARIES

**Anonymous authors**

Paper under double-blind review

## ABSTRACT

We introduce the Solid Harmonic Wavelet Bispectrum, an operator for 2D images that computes third-order correlations over angular frequency components of solid harmonic wavelet responses. By using angular rather than spatial frequencies, our bispectrum achieves lower dimensionality than traditional 2D scattering-based bispectra, avoiding comparisons across two spatial dimensions while still preserving rich frequency information. Extending these bispectra to first- and second-order scattering coefficients produces low-dimensional multi-scale features that capture detailed image structure. To illustrate the quality of the representations, we use k-nearest neighbors, which highlights that our features encode meaningful similarity structure even without a learned parametric classifier. Results on texture, medical, and galaxy images demonstrate that these features show improved separability and similarity structure compared to existing geometric and deep learning-based representations.

## 1 INTRODUCTION

Designing discriminative features that remain robust to geometric variability is essential in supervised learning. A common strategy is to exploit symmetries—transformations such as translations and rotations that preserve labels—by constructing representations that are covariant under these actions and then pooling to obtain invariants. This principle underlies many methods in harmonic analysis and group-based learning, and is particularly valuable in domains like computational chemistry, astronomy, and medical imaging, where data is costly and interpretability matters.

Convolutional Neural Networks (CNNs) achieve strong performance through end-to-end optimization, but they lack built-in geometric priors and typically require large datasets and augmentation. Geometric deep learning introduces equivariant architectures to encode symmetry, yet these approaches depend on predefined groups, large parametric models, and heavy training, which can reduce transparency and compromise analytical guarantees. By contrast, wavelet scattering networks (Bruna & Mallat, 2013) provide invariance by design using cascades of predefined filters, modulus nonlinearities, and averaging. Wavelet Scattering operators have matched deep learning baselines with simple classifiers in tasks such as texture discrimination (Sifre & Mallat, 2014), quantum molecular regression (Eickenberg et al., 2017), and object classification (Oyallon et al., 2014). A key limitation, however, is that the modulus operator discards phase information, restricting scattering networks to second-order statistics.

We introduce the *Solid Harmonic Wavelet Bispectrum* (SHWB), an operator that unifies multi-scale analysis, higher-order statistics, and group invariance. Bispectral analysis preserves phase couplings lost by power spectra or modulus operators, but prior formulations in the Fourier domain were not adapted to geometric invariance. Our approach computes third-order correlations over solid harmonic wavelet responses, capturing nonlinear interactions between rotational components at each scale. By spanning one angular frequency instead of two spatial frequencies, SHWB yields compact, frequency-preserving representations that are covariant to roto-translations and can be integrated to form invariants. Extending bispectral analysis to first- and second-order scattering coefficients further enriches the representation while keeping dimensionality low.

054 A key methodological choice is to evaluate SHWB features in isolation with simple classifiers, fol-  
055 lowing prior work on scattering (Bruna & Mallat, 2013; Eickenberg et al., 2017) and linear probes for  
056 representation learning (Alain & Bengio, 2016). This avoids conflating the quality of the representa-  
057 tion with classifier capacity, reduces computational cost, and highlights the intrinsic discriminative  
058 power of the features. Despite this minimalist evaluation, SHWB achieves competitive or superior  
059 performance to deep learning baselines across textures, medical images, and a  $\sim 20\text{K}$ -image astro-  
060 physics dataset, where simple models such as k-nearest neighbors already reveal strong similarity  
061 structure in the proposed feature space.

## 062 063 064 065 066 2 RELATED WORK

067  
068  
069 **Wavelet and Scattering Methods.** Invariant feature extraction has long been studied in harmonic  
070 analysis. Multiresolution analysis (Mallat, 1989) and steerable wavelets (Freeman & Adelson, 1990;  
071 Simoncelli et al., 1992) decompose signals into multi-scale orientation-separable filters, providing  
072 translation and rotation invariance. Wavelet scattering networks (Bruna & Mallat, 2013) extend this  
073 with cascades of wavelet transforms, modulus nonlinearities, and spatial averaging, offering stabi-  
074 lity to deformations while retaining interpretability. Steerable (Sifre & Mallat, 2013) and solid  
075 harmonic wavelets (Eickenberg et al., 2017) further capture roto-translation invariants, achieving  
076 strong results in texture classification, molecular regression, and related tasks. The main draw-  
077 back is that the modulus operator discards phase and restricts scattering to second-order statistics,  
078 although Mallat et al. (2019) showed how phase correlations can be reintroduced in sparse signals.  
079 Recent work on scattering spectra Cheng et al. (2024) extends scattering to higher-order statistics for  
080 generative modeling of physical systems, using maximum entropy models and covariances of scat-  
081 tering coefficients to generate new field realizations. Our work extends this line by incorporating  
082 bispectral statistics into solid harmonic scattering to preserve discriminative phase interactions for  
083 supervised learning, computing bispectral coefficients explicitly while maintaining analytical rota-  
084 tion invariance and interpretability. Other recent developments include Rodriguez et al. (2019), who  
085 augmented scattering with learned CNN features for image classification, and Saydjari & Finkbeiner  
086 (2022), who developed efficient computational algorithms for scattering transforms. Our approach  
087 differs by focusing on analytical bispectral features without learned parameters.

088 **Geometric Deep Learning.** Geometric deep learning develops neural architectures that respect the  
089 symmetries of data by ensuring equivariance to group actions. Group-equivariant CNNs (G-CNNs)  
090 (Cohen & Welling, 2016) and harmonic networks (H-Nets) (Worrall et al., 2017) exemplify this  
091 approach, where feature maps transform predictably under rotations or translations. Invariance is  
092 typically obtained only after additional pooling or integration over the group. Bispectral neural  
093 networks (Sanborn et al., 2023) move closer to invariance by learning bispectral coefficients that  
094 collapse transformed inputs into identical representations, but this requires large parametric models  
095 and orbit-based supervision. By contrast, our method produces invariants analytically and by con-  
096 struction, without training or group-specific pooling, allowing simple classifiers to operate directly  
097 on the features while retaining interpretability and computational efficiency.

098 **Higher-Order Spectra.** Higher-order statistics, such as the bispectrum, capture phase correla-  
099 tions lost by second-order statistics. Higher-order spectral analysis originates from studies of non-  
100 Gaussian signals, with applications in fields such as oceanography, telecommunications, and plasma  
101 physics (Nikias & Raghuvver, 1987). This was extended to multi-scale bispectral analysis over  
102 wavelet transforms (van Milligen et al., 1995; Jamšek et al., 2007; Jamšek et al., 2010), while New-  
103 man et al. (2021) proposed a normalised "bispectral density" subject to wavelet parametrisation.  
104 However, such approaches remain sensitive to rotation, lacking geometric invariance. On the other  
105 hand, Kakarala (2012) projected signals onto solid harmonics before computing the Fourier bis-  
106 pectrum. Kondor (2007) projected images onto the unit sphere and decomposed them into solid  
107 harmonics, though at the cost of distortions due to spherical projection and redundant computation.

While prior work has addressed invariance or higher-order statistics separately, few methods unify  
both in an interpretable, efficient, and empirically competitive manner. SHWB aims to fill this gap.

### 3 SOLID HARMONIC WAVELET BISPECTRUM

#### 3.1 BACKGROUND: SOLID HARMONIC WAVELET TRANSFORM

The Solid Harmonic Wavelet Transform (SHWT) Eickenberg et al. (2017) provides multiscale roto-translation covariant features for 2D and 3D signals. Solid harmonics, solutions of the Laplace equation in polar or spherical coordinates, form a complete basis on the unit circle or sphere. Localized solid harmonic wavelets are obtained by modulating them with a Gaussian envelope, yielding band-pass, zero-mean filters.

In 2D, a wavelet at scale  $j$  and angular frequency  $\ell$  is

$$\psi_{j,\ell}(r, \varphi) = \frac{1}{\sqrt{(2\pi\sigma_j^2)^2}} e^{-r^2/2\sigma_j^2} r^\ell e^{i\ell\varphi}, \quad \sigma_j = 2^j\sigma_0, \quad (1)$$

normalized to unit energy. Varying  $j$  and  $\ell$  produces a filter bank; the case  $\ell = 0$  reduces to an isotropic Gaussian low-pass. This aligns with the principles of multiresolution analysis (Mallat, 1989), enabling the capture of both coarse- and fine-scale structures. A visualization of the resulting wavelets is included in Appendix A.2.

For a signal  $x$  and location  $u = (r \cos \varphi, r \sin \varphi)^T$ , the solid harmonic wavelet transform at scale  $j$  and frequency  $\ell$  is

$$W[x, j, \ell](u) = (x \star \psi_{j,\ell})(u). \quad (2)$$

The Solid Harmonic Wavelet Scattering transform applies a modulus non-linearity to remove phase and yield roto-translation covariance,

$$U[x, j, \ell](u) = |x \star \psi_{j,\ell}|(u). \quad (3)$$

Higher-order features are obtained by cascading this operation, e.g.

$$U[x, j_1, \ell_1, j_2, \ell_2](u) = |U[x, j_1, \ell_1] \star \psi_{j_2, \ell_2}|(u). \quad (4)$$

Integration over  $u$  reduces these covariant responses to invariants spanning scales  $j_*$  and angular frequencies  $\ell_*$ . However, the modulus discards phase information, which motivates redundant filter-bank designs and dedicated phase-recovery methods Waldspurger et al. (2015). In contrast, higher-order spectra explicitly reintroduce phase relationships by encoding nonlinear interactions between frequency components. We review these next.

#### 3.2 BACKGROUND: HIGHER-ORDER SPECTRA

Second-order statistics, such as the power spectrum  $P(f) = |\hat{x}(f)|^2$ , describe the distribution of energy across frequencies but discard phase information, limiting their ability to represent structure in natural signals (Oppenheim & Lim, 1981; Skarbnik et al., 2010).

Higher-order spectra extend to third- and fourth-order statistics, namely the bispectrum and trispectrum. The bispectrum captures quadratic phase coupling among components  $f_1$ ,  $f_2$ , and  $f_1 + f_2$ :

$$B(f_1, f_2) = \hat{x}(f_1) \hat{x}(f_2) \overline{\hat{x}(f_1 + f_2)}. \quad (5)$$

If we write the above terms as complex numbers of amplitude  $a_f$  and phase  $\varphi_f$ , we have  $a_{f_1} a_{f_2} a_{f_1+f_2} e^{i(\varphi_{f_1} + \varphi_{f_2} - \varphi_{f_1+f_2})}$ . Therefore, the bispectrum maintains information about the phase interactions but its magnitude also reflects underlying power at each frequency. To separate coupling from power, one defines the *bicoherence*:

$$\text{Bic}(f_1, f_2) = \frac{|B(f_1, f_2)|}{\sqrt{|\hat{x}(f_1)|^2 |\hat{x}(f_2)|^2 |\hat{x}(f_1 + f_2)|^2}}, \quad (6)$$

which is bounded between 0 and 1, with values near 1 indicating strong phase coupling.

### 3.3 SOLID HARMONIC WAVELET BISPECTRUM

The Solid Harmonic Wavelet Bispectrum (SHWB) extends the scattering framework by capturing higher-order, phase-preserving interactions between solid harmonic wavelet responses across angular frequencies. For a signal  $x$ , at scale  $j$  and angular frequencies  $\ell_1, \ell_2$ , we define the bispectrum in the Fourier domain as

$$\text{SHWB}[x, j, \ell_1, \ell_2](f) = \hat{W}[x, j, \ell_1](f) \cdot \hat{W}[x, j, \ell_2](f) \cdot \overline{\hat{W}[x, j, \ell_1 + \ell_2](f)}, \quad (7)$$

where  $f$  indexes spatial frequencies (horizontal and vertical), and  $\hat{W}[x, j, \ell]$  is the Fourier transform of the wavelet response.

As shown by Eickenberg et al. (2017), a rotation  $R_\gamma \in \text{SO}(2)$  of angle  $\gamma$  applied to a signal  $x$  acts on the solid harmonic wavelet responses by a multiplicative phase factor:

$$W[R_\gamma x, j, \ell] = e^{i\ell\gamma} R_\gamma W[x, j, \ell]. \quad (8)$$

Substituting Equation 8 into 7, and using the constraint  $\ell_3 = \ell_1 + \ell_2$ , the phase factors cancel, yielding rotational covariance:

$$\text{SHWB}[R_\gamma x, j, \ell_1, \ell_2](f) = R_\gamma \text{SHWB}[x, j, \ell_1, \ell_2](f). \quad (9)$$

so the bispectrum is rotation-covariant. In contrast to scattering networks with Morlet wavelets, which achieve invariance by discarding phase via a modulus nonlinearity, SHWB **retains complex phase** information while still achieving covariance.

**Roto-translation invariance and  $L_p$  integration.** Integration over spatial frequencies produces roto-translation invariance: while rotation of the input  $x$  rotates the representation in the frequency domain, the integral over all frequencies remains unchanged. For real-valued signals, Hermitian symmetry ( $\hat{x}(-f) = \overline{\hat{x}(f)}$ ) ensures that this integral yields a real-valued scalar:

$$\text{SHWB}[x, j, \ell_1, \ell_2] = \int_f \text{SHWB}[x, j, \ell_1, \ell_2](f) df. \quad (10)$$

To enrich the representation and capture different aspects of the spatial-frequency distribution, we compute these integrals using various  $L_p$  norms:

$$\text{SHWB}[x, j, \ell_1, \ell_2, p] = \left( \sum_f |\text{SHWB}[x, j, \ell_1, \ell_2](f)|^p \right)^{\frac{1}{p}}, \quad (11)$$

Varying  $p$  provides estimates of different moments of the bispectral distribution, enhancing sensitivity to the structure of nonlinear feature interactions.

**Bicoherence.** Normalizing by the magnitude of each term isolates the strength of phase alignment independently of signal amplitude:

$$\text{SHWBic}[x, j, \ell_1, \ell_2] = \frac{|\text{SHWB}[x, j, \ell_1, \ell_2]|}{\sqrt{\int_f |\hat{W}[x, j, \ell_1](f)|^2 df \int_f |\hat{W}[x, j, \ell_2](f)|^2 df \int_f |\hat{W}[x, j, \ell_1 + \ell_2](f)|^2 df}} \quad (12)$$

which takes values in  $[0, 1]$  with 1 indicating strong phase coupling between  $\ell_1$  and  $\ell_2$ .

#### 3.3.1 COMMENTS ON COMPUTATIONAL EFFICIENCY

A critical observation is that by computing the bispectrum over angular frequencies rather than Cartesian frequency pairs, SHWB efficiently spans the 2D space of nonlinear interactions with coefficients that scale quadratically with the number of angular frequencies as opposed to quartic scaling in the spatial dimensions.

Additionally, to reduce computational cost, we apply the following strategies:

- **Precomputation:** Wavelet responses  $W[x, j, \ell]$  are computed once and cached, reused across multiple bispectral terms. Mini-batching manages memory for large datasets.
- **Symmetry:** SHWB is symmetric in  $\ell_1$  and  $\ell_2$ , so we compute only  $\ell_1 < \ell_2$ .
- **Acceleration:** Computations are implemented in PyTorch with FFT-based convolutions, leveraging GPU acceleration and the Kymatio framework Andreux et al. (2018).

### 3.4 SOLID HARMONIC WAVELET SCATTERING BISPECTRUM

The scattering transform, and solid harmonic wavelet scattering by extension, attempt to obtain high order information by a cascade of wavelets in a manner similar to Deep Learning. In this section, we demonstrate that this approach, while proposed as an alternative to higher order spectra, can easily be combined with higher order spectra to provide rich feature spaces. We do that by demonstrating that the bispectrum can be used as an alternative non-linearity at the last layer/order of a scattering transform, while maintaining a cascade of wavelet convolutions and modulus non-linearities for intermediate layers.

As in relevant literature (Mallat, 2012; Eickenberg et al., 2017), we compute the bispectrum over first- and second- order scattering coefficients. The zeroth-order bispectral coefficient is computed as in other scattering transforms using the signal integral, or in practice by calculating  $L_p$  to capture different proxies for the raw moments of the distribution of the data:

$$\text{SHWSB}_0[x, p] = \|x\|_p \quad (13)$$

For the first and second order coefficients, we introduce the following notation for the linear responses of the first- and second-order scattering:

$$W[x, j_1, \ell](u) = x \star \psi_{j_1, \ell}(u) \quad (14)$$

$$W[x, j_1, j_2, \ell](u) = |x \star \psi_{j_1, \ell}| \star \psi_{j_2, \ell}(u) \quad (15)$$

For all  $\ell_1, \ell_2 \in L$  with  $\ell_1 < \ell_2$ , and  $\ell_1 + \ell_2 = \ell_3 \in L$ , we define the bispectral coefficients:

$$\text{SHWSB}[x, j_1, \ell_1, \ell_2](f) = \hat{W}[x, j_1, \ell_1](f) \cdot \hat{W}[x, j_1, \ell_2](f) \cdot \overline{\hat{W}[x, j_1, \ell_3](f)} \quad (16)$$

$$\text{SHWSB}[x, j_1, j_2, \ell_1, \ell_2](f) = \hat{W}[x, j_1, j_2, \ell_1](f) \cdot \hat{W}[x, j_1, j_2, \ell_2](f) \cdot \overline{\hat{W}[x, j_1, j_2, \ell_3](f)}, \quad (17)$$

for first and second order coefficients. The first order invariant,  $\text{SHWSB}[x, j_1, \ell_1, \ell_2]$ , and second order invariant,  $\text{SHWSB}[x, j_1, j_2, \ell_1, \ell_2]$ , are obtained through integration over  $f$  in Equations 16 and 17 respectively. Before integration, the coefficients remain complex-valued, encoding relative phase information across angular frequencies; after integration, Hermitian symmetry ensures that, for real signals, the resulting invariants are real-valued. The bicoherence ( $\text{SHWSBic}$ ) is again derived by normalising with the product of the scattering magnitudes:

$$\text{SHWSBic}[x, j_1, \ell_1, \ell_2] = \frac{\text{SHWSB}[x, j_1, \ell_1, \ell_2]}{\|\hat{W}[x, j_1, \ell_1]\| \cdot \|\hat{W}[x, j_1, \ell_2]\| \cdot \|\hat{W}[x, j_1, \ell_3]\|} \quad (18)$$

$$\text{SHWSBic}[x, j_1, j_2, \ell_1, \ell_2] = \frac{\text{SHWSB}[x, j_1, j_2, \ell_1, \ell_2]}{\|\hat{W}[x, j_1, j_2, \ell_1]\| \cdot \|\hat{W}[x, j_1, j_2, \ell_2]\| \cdot \|\hat{W}[x, j_1, j_2, \ell_3]\|} \quad (19)$$

As with SHWB, the SHWSB transform can perform the integration over spatial frequencies for Equations 16–17 by applying  $L_p$  pooling on the spatial coordinates of the representations, which we denote with the parameter  $p$  as:

$$\text{SHWSB}[x, j_1, \ell_1, \ell_2, p] = \left( \sum_f |\text{SHWSB}[x, j_1, \ell_1, \ell_2](f)|^p \right)^{\frac{1}{p}} \quad (20)$$

When  $L_p$  pooling is applied, the normalization in Equations 18 and 19 is adjusted by replacing raw magnitudes with their pooled values, ensuring that  $\text{SHWSBic}$  remains bounded in  $[0, 1]$ .

First-order bispectral coefficients are equivalent to our definition of SHWB, and so the property in Equation 9 holds - they are roto-translation covariant by definition, and reduce to invariants through  $L_p$  pooling. The formal proof of roto-translation covariance is detailed in Appendix A.3.

Second-order scattering coefficients retain the modulus nonlinearity after the first convolution, preserving stability. Although the modulus discards inter-scale phase, we argue that this phase loss is offset by increased robustness, consistent with scattering theory. Phase interactions within the same scale are retained through the bispectral term. Under rotation, a signal’s convolution with a solid harmonic wavelet transforms by a phase factor, as shown in Equation 8. The modulus operator applied on the right side of Equation 8 removes this factor, leaving:

$$U[R_\gamma x, j, \ell](u) = R_\gamma(x \star \psi_{j, \ell})(u) \quad (21)$$

Since solid harmonics commute with rotation, a second convolution preserves rotation structure. Therefore, the same property as in Equation 9 establishes that both first- and second-order SHWSB coefficients are rotation covariant.  $L_p$  pooling converts them to invariants.

## 4 EXPERIMENTS

We benchmark using simple models, parametrising  $\ell$  over a coarse sweep and selecting  $j$  relative to the scale of features relevant to the task. In each experiment, we report the best result for each representation, avoiding arbitrary parameter choices that best suit our new operators. The main strategy of the experiments is to demonstrate the ability of the Solid Harmonic Transforms to identify novel and interesting image features. Wavelet-based representations have consistently achieved comparable results to deep learning (Eickenberg et al., 2017; Zarka et al., 2020) with simple classifiers. More complex architectures, combining scattering with large neural networks (Patro & Agneswaran, 2023), require an exorbitant amount of resources while offering little understanding of the representation.

### 4.1 TEXTURE CLASSIFICATION

Textural similarity is primarily attributed to Fourier phase, rather than magnitude, information with respect to human perception (Dong et al., 2017), making its preservation crucial in texture analysis. The KTH-TIPS benchmark (Fritz et al., 2004) remains a relevant testbed for examining the impact of higher-order, phase-preserving operators, particularly in domains where structural information drives variability—interactions encoded by the Solid Harmonic Wavelet Bispectrum. The dataset consists of 10 classes, each with 81 samples varying in scale, shear, and illumination.

We test our representations over 200 random train-test splits for different training set sizes, as done in (Sifre & Mallat, 2013). We train linear SVMs on our representations, which hold similar properties to the PCA classifier used in (Sifre & Mallat, 2013) over translation invariant (Mallat, 2012) and roto-translation invariant scattering (Sifre & Mallat, 2013). Enhanced roto-translation scattering refers to roto-translation scattering followed by a logarithmic nonlinearity, averaging over scales, and scale augmentations. Therefore, the most direct comparisons are between translation scattering, roto-translation scattering, and our proposed operators.

Table 1: Accuracy and standard deviation over 200 random training subsets with varying sizes of training set on KTH-TIPS

Features	5 samples	20 samples	40 samples
Translation scattering	69.1 $\pm$ 3.5	94.8 $\pm$ 1.3	98.0 $\pm$ 0.8
Roto-trans scattering	69.5 $\pm$ 3.6	94.9 $\pm$ 1.4	98.3 $\pm$ 0.9
SHWS	86.9 $\pm$ 0.8	96.1 $\pm$ 0.3	98.5 $\pm$ 0.4
SHWB	87.7 $\pm$ 1.7	94.8 $\pm$ 0.5	96.9 $\pm$ 0.3
SHWBic	87.2 $\pm$ 1.3	96.5 $\pm$ 0.6	99.3 $\pm$ 0.3
SHWSB	<b>88.6 <math>\pm</math> 2.5</b>	95.1 $\pm$ 0.3	97.4 $\pm$ 0.3
SHWSBic	85.1 $\pm$ 0.9	97.0 $\pm$ 0.7	98.6 $\pm$ 0.2
Enhanced roto-trans	84.3 $\pm$ 3.1	<b>98.3 <math>\pm</math> 0.9</b>	<b>99.4 <math>\pm</math> 0.4</b>

SHWBic and SHWSBic outperform previous isolated scattering representations across all training sizes, demonstrating the added expressivity of bispectral representations in capturing higher-order nonlinearities. The improvement is most pronounced for the smallest training set ( $n = 5$ ), where SHWSB achieves 88.6% accuracy—an increase of almost 20 points compared to translation scattering. In this low-data regime, the bispectrum outperforms corresponding bicoherences, albeit with larger uncertainty (e.g.,  $\pm 2.5$  for SHWSB), suggesting that the additional information on total energy in frequency couplings introduces meaningful variability for classification.

As training sizes increase, this relationship shifts: bicoherence-based representations (SHWBic and SHWSBic) improve considerably, with 20-sample and 40-sample results reaching 96.5% and 99.3% accuracy, respectively, highlighting that patterns in the strength of phase coupling become more informative than raw energy. These results underscore the crucial role of phase-preserving, higher-order interactions for tasks driven by structural variability.

Solid Harmonic Wavelets consistently outperform Morlet scattering, emphasizing their capacity to encode textural information. While enhanced roto-translation scattering reaches high accuracy in larger training sets, SHWSBic achieves comparable performance with more certainty in intermediate

sizes, demonstrating the robustness of our operators across data regimes. Overall, these experiments highlight how the combination of bispectral interactions and roto-translation invariance produces rich, discriminative features that reduce uncertainty relative to earlier scattering-based approaches.

## 4.2 MEDICAL IMAGING

Deep learning in medical imaging is often limited by the availability of large-scale datasets (Zhou et al., 2021). Symmetry-aware wavelet operators mathematically impose key structural invariances, simplifying the variability within the representation and yielding meaningful features without large-scale optimisation. We evaluate this on MedMNIST (Yang et al., 2021; 2023), focusing on  $28 \times 28$  images for fair comparison with existing deep learning baselines. MedMNIST presents distinct challenges: BreastMNIST requires detection of subtle shape irregularities, RetinaMNIST demands small vascular anomaly identification, and DermaMNIST relies on complex textural patterns—all benefiting from rotation-invariant analysis.

**BreastMNIST (780 samples)** is a binary classification task distinguishing benign from malignant tumors, which are discriminable based on symmetry and shape (Wei et al., 2020).

**RetinaMNIST (1,600 samples)** is an ordinal regression task grading diabetic retinopathy severity from 1–5, based on subtle vascular features (Porwal et al., 2020).

**DermaMNIST (10,015 samples)** is a seven-class classification task of pigmented skin lesions distinguished by asymmetry, border regularity, and color distribution.

For BreastMNIST, we use larger scales ( $J \in [3, 4]$ ) to capture global symmetries. For RetinaMNIST and DermaMNIST, we focus on  $J = 2$  for localized feature capture. Linear classifiers with  $L_2$  regularization and singular value decomposition are used to isolate feature effectiveness.

Table 2: Accuracy over MedMNIST benchmark datasets. We performed an initial hyperparameter sweep of H-Net for each task over ring discretizations, number of filters, and angular frequencies. Other baselines are from Yang et al. (2021; 2023).

Model	BreastMNIST	RetinaMNIST	DermaMNIST
ResNet-18 (Yang et al., 2021; 2023)	86.3%	52.5%	73.5%
ResNet-50 (Yang et al., 2021; 2023)	81.2%	52.8%	73.5%
auto-sklearn (Yang et al., 2021; 2023)	80.3%	51.5%	71.9%
AutoKeras (Yang et al., 2021; 2023)	83.1%	50.3%	74.9%
Google AutoML (Yang et al., 2021; 2023)	86.1%	53.1%	<b>76.8%</b>
H-Net (Worrall et al., 2017)	84.4%	48.7%	72.4%
SHWS	84.6%	53.5%	72.1%
SHWB	78.2%	52.8%	69.7%
SHWBic	82.7%	<b>54.5%</b>	71.7%
SHWSB	85.9%	54.3%	71.0%
SHWSBic	<b>86.5%</b>	51.0%	72.2%

In data-scarce tasks (BreastMNIST, RetinaMNIST), our approaches surpass most baselines. SHWS-Bic achieves 86.5% accuracy on BreastMNIST, outperforming both ResNet-18 (86.3%), and H-Nets (84.4%). Our methods are particularly effective for medical images, which often contain Gaussian noise (to which the bispectrum is inherently insensitive), variable orientations (addressed by our roto-translation invariance), and multiscale features (captured by our wavelet decomposition). With larger datasets (DermaMNIST), our fixed-filter approaches remain competitive. Linear separability is evident: for instance, a Gaussian SVM on SHWSBic achieves 75.9% on DermaMNIST, outperforming all but one baseline.

SHWSBic outperforms the SHWBic on BreastMNIST, suggesting that a more sophisticated set of features is required to capture the variability of malignant tumors. Interference across scales is particularly important for that dataset. However, in RetinaMNIST key indicators are blurry blob-like structures with localised shape, making cross-scale interactions less discriminative. In this case, the scattering hierarchy in SHWSBic introduces features to minimal effect and SHWBic achieves the highest accuracy. Overall, these experiments highlight the advantage of built-in geometric invariances for data-efficient architectures.

### 4.3 ASTROPHYSICS

We study the application of SHWSBic in astrophysics. Galaxy mergers show complex features that have historically been challenging to identify using simple statistical measures (Pawlik et al., 2016; Conselice, 1997). Galaxy mergers exhibit complex, extended morphological structures that evolve over time and vary with mass ratio – precisely the type of multiscale geometric information our methods intend to capture.

We use a sample of galaxy mergers with mass ratios  $\mu \geq -2$  within 2.5 Gyrs of a merger for a redshift up to 0.1 (following methodology of Avirett-Mackenzie et al. (2024)) derived from IllustrisTNG (Nelson et al., 2019). Data are extracted to match high spatial resolution from e.g. Hubble Space Telescope and do not have observational noise added. We downsample the initial dataset into two separate datasets for the regression of both time since last merger ( $n=18,791$ ) and log merger mass ratio ( $n=16,870$ ). Log merger mass ratio has mean  $-1.10$  and standard deviation  $0.54$ , while time since last merger has mean  $1.93$  Gyr and standard deviation  $1.06$  Gyr. Images are cropped to  $32\times$  the half stellar mass radius (as calculated from Nelson et al. (2019)) and resized to a resolution of  $128\times 128$ . We optimise subject to mean-absolute-error (MAE). Experiments use  $j = 4$  to capture larger, broader symmetries such as tidal tails.

Table 3: Results of merger feature regression on  $128 \times 128$  resolution samples.

Model	Log $\mu$		$\Delta t$ (Gyr)	
	MAE	RMSE	MAE	RMSE
Random predictor	0.6479	0.7932	1.2306	1.5065
Oriented scattering + linear	0.4119	0.4953	0.8025	0.9604
SHWS + linear	0.3876	0.4701	0.7781	0.9360
SHWS + KNN	0.3539	0.4699	0.7264	0.9370
<b>SHWSBic + KNN</b>	<b>0.3322</b>	<b>0.4498</b>	<b>0.6633</b>	0.9139
CNN	0.3564	0.4586	0.7206	0.9406
H-Net	0.3581	0.4543	0.7345	<b>0.9015</b>

Table 3 shows that models trained on our operators effectively regress merger features. Best performing models achieve MAE of two-thirds the standard deviation for each target. We compare our approaches to a baseline random predictor, where predictions are sampled from the target range - demonstrating an almost 50% reduction in error rates. We observe an improvement using SHWS over oriented Morlet wavelets when input to linear models. We find significant improvements by training non-parametric models on our coefficients (KNN regressors with inverse  $L_1$  distance weighting). This suggests that accurate prediction of merger features requires capturing not just global symmetries but also complex relationships between features at varying scales—precisely what SHWSBic encodes. The non-parametric approach enables the model to learn additional nonlinearities from our coefficients without imposing restrictive functional forms. SHWSBic achieves the best MAE across both tasks, though H-Nets obtain lower RMSE for time prediction. The difference between MAE and RMSE performance suggests that while our method achieves better typical-case accuracy, H-Nets may be more robust to outliers—possibly due to learned features adapting to rare morphological configurations. This highlights a characteristic trade-off: analytical features excel at capturing common structural patterns but may be less flexible for atypical cases without data-driven adaptation. Figure 1 shows true-predicted plots.

While previous studies have focused on binary merger classification (Avirett-Mackenzie et al., 2024; Walmsley et al., 2024; de Graaff et al., 2025), our approach pioneers regression on continuous merger features from morphological data. This is particularly valuable in astrophysics, where quantifying merger parameters typically requires extensive spectroscopic follow-up or access to simulation data. The strong relationship between true and predicted values (Figure 1) demonstrates that our wavelet-based approach effectively captures the complex structural signatures of merger dynamics, suggesting applications to large-scale observational surveys.

Figure 2 demonstrates the scaling of the predictive precision with respect to dataset size. All algorithms approach a similar MAE with increasing data points. The interesting observation is on

432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485

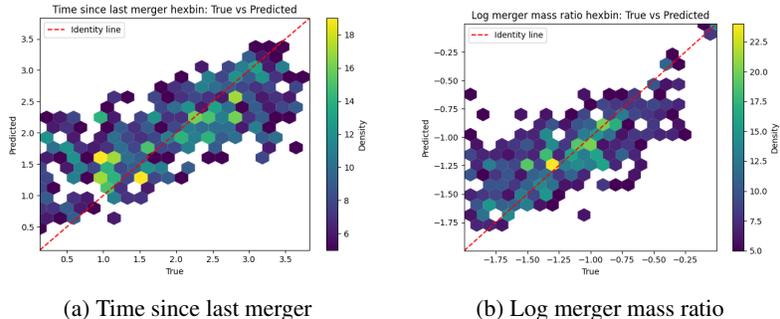


Figure 1: True-predicted hexbin plots using K-Nearest-Neighbours regressor with SHWSBic input. Hexbins with more than 5 samples are shown to highlight underlying trends and remove outliers.

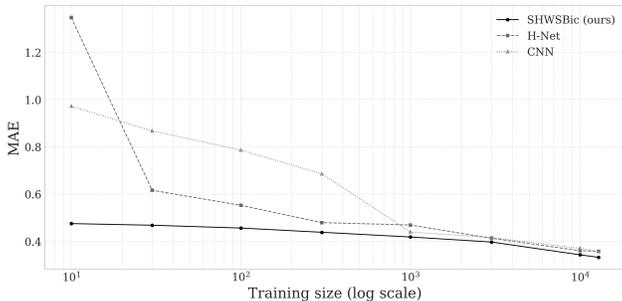


Figure 2: Comparison of Mean Absolute Error for log mass ratio regression with increasing training set sizes (log scale). SHWSBic features are used with a KNN regressor.

the low data regime, where our method substantially outperforms Deep Learning and Geometric Deep Learning baselines. The low-data regime is interesting due to the rarity of galaxy mergers in both simulations and even more in real observational data. Additionally, Figure 2 supports the case that the SHWSBic space is a natural space for comparisons of galaxies. Hybrid approaches combining wavelet features with trainable layers could further improve performance (Zarka et al., 2020; Oyallon et al., 2014); however, naively minimizing errors could obscure the contribution of the representation relative to the supervised learning algorithm.

## 5 CONCLUSION

We introduce new operators that generate roto-translation invariant features for supervised learning over image-based tasks. The **Solid Harmonic Wavelet Bispectrum (SHWB)** captures multi-scale interactions between feature types, while preserving phase information. Building on this, the **Solid Harmonic Wavelet Scattering Bispectrum (SHWSB)** combines properties of both scattering networks and the solid harmonic wavelet bispectrum, enabling the encoding of higher-order nonlinearities while retaining critical phase information.

Experiments with simple linear models demonstrate that these representations are increasingly discriminative and perform competitively with deep learning in limited-data settings. For textures, SHWB and SHWSB capture structural variability and higher-order interactions, outperforming standard scattering representations. In medical imaging, our operators effectively linearize decision boundaries and achieve state-of-the-art or competitive accuracy on small datasets. In astrophysics, SHWSBic successfully predicts merger features, particularly in low-data regimes, highlighting the natural expressivity of phase-preserving, multi-scale operators.

A limitation of our study is the lack of evaluation on large-scale datasets such as ImageNet. While such benchmarks are standard for deep learning, they may not accurately reflect the advantages of our operators. Hybrid approaches combining wavelet features with trainable layers (Oyallon et al.,

2014; Patro & Agneeswaran, 2023) could further enhance performance, but their inclusion would obscure the contribution of the operator itself and add significant computational overhead.

Overall, our work highlights the crucial role of phase-preserving, higher-order interactions in extracting discriminative features. The proposed operators provide interpretable, mathematically grounded, and data-efficient representations well-suited for tasks where symmetry and multi-scale structure are central.

## REFERENCES

- Guillaume Alain and Yoshua Bengio. Understanding intermediate layers using linear classifier probes. *arXiv preprint arXiv:1610.01644*, 2016.
- Mathieu Andreux, Tomás Angles, Georgios Exarchakis, Roberto Leonarduzzi, Gaspar Rochette, Louis Thiry, John Zarka, Stéphane Mallat, Joakim Andén, Eugene Belilovsky, Joan Bruna, Vincent Lostanlen, Matthew J. Hirn, Edouard Oyallon, Sixin Zhang, Carmine-Emanuele Cella, and Michael Eickenberg. Kymatio: Scattering transforms in python. *CoRR*, abs/1812.11214, 2018. URL <http://arxiv.org/abs/1812.11214>.
- M. S. Avirett-Mackenzie, C. Villforth, M. Huertas-Company, S. Wuyts, D. M. Alexander, S. Bonoli, A. Lapi, I. E. Lopez, C. Ramos Almeida, and F. Shankar. A post-merger enhancement only in star-forming Type 2 Seyfert galaxies: the deep learning view. *Monthly Notices of the Royal Astronomical Society*, 528(4):6915–6933, March 2024. doi: 10.1093/mnras/stae183.
- Joan Bruna and Stephane Mallat. Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1872–1886, 2013. doi: 10.1109/TPAMI.2012.230.
- Sihao Cheng, Rudy Morel, Erwan Allys, Brice Ménard, and Stéphane Mallat. Scattering spectra models for physics. *PNAS Nexus*, 3(4):pgae103, 03 2024. ISSN 2752-6542. doi: 10.1093/pnasnexus/pgae103. URL <https://doi.org/10.1093/pnasnexus/pgae103>.
- Taco Cohen and Max Welling. Group equivariant convolutional networks. In Maria Florina Balcan and Kilian Q. Weinberger (eds.), *Proceedings of The 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 2990–2999, New York, New York, USA, 20–22 Jun 2016. PMLR. URL <https://proceedings.mlr.press/v48/cohen16.html>.
- C. J. Conselice. The Symmetry, Color, and Morphology of Galaxies. *Publications of the Astronomical Society of the Pacific*, 109:1251–1255, November 1997. doi: 10.1086/134004.
- Rosa de Graaff, Berta Margalef-Bentabol, Lingyu Wang, Antonio La Marca, William J. Pearson, Vicente Rodriguez-Gomez, and Mike Walmsley. Classifying merger stages with adaptive deep learning and cosmological hydrodynamical simulations. *arXiv e-prints*, art. arXiv:2502.06254, February 2025. doi: 10.48550/arXiv.2502.06254.
- Xinghui Dong, Ying Gao, Junyu Dong, and Mike J. Chantier. The importance of phase to texture similarity. In *2017 IEEE International Conference on Computer Vision Workshops (ICCVW)*, pp. 2758–2766, 2017. doi: 10.1109/ICCVW.2017.324.
- Michael Eickenberg, Georgios Exarchakis, Matthew Hirn, and Stéphane Mallat. Solid harmonic wavelet scattering: Predicting quantum molecular energy from invariant descriptors of 3d electronic densities. In *Advances in Neural Information Processing Systems*, pp. 6540–6549, 2017.
- W.T. Freeman and E.H. Adelson. Steerable filters for early vision, image analysis, and wavelet decomposition. In *Proceedings Third International Conference on Computer Vision*, pp. 406,407,408,409,410,411,412,413,414,415, Los Alamitos, CA, USA, December 1990. IEEE Computer Society. doi: 10.1109/ICCV.1990.139562. URL <https://doi.ieeeecomputersociety.org/10.1109/ICCV.1990.139562>.
- Mario Fritz, Eric Hayman, Barbara Caputo, and Jan-Olof Eklundh. The kth-tips database. 2004.

- 540 Janez Jamšek, Milan Paluš, and Aneta Stefanovska. Detecting couplings between interacting oscillators with time-varying basic frequencies: Instantaneous wavelet bispectrum and information  
541 theoretic approach. *Phys. Rev. E*, 81:036207, Mar 2010. doi: 10.1103/PhysRevE.81.036207.  
542 URL <https://link.aps.org/doi/10.1103/PhysRevE.81.036207>.  
543  
544
- 545 Janez Jamšek, Aneta Stefanovska, and Peter V. E. McClintock. Wavelet bispectral analysis for the  
546 study of interactions among oscillators whose basic frequencies are significantly time variable.  
547 *Phys. Rev. E*, 76:046221, Oct 2007. doi: 10.1103/PhysRevE.76.046221. URL <https://link.aps.org/doi/10.1103/PhysRevE.76.046221>.  
548
- 549 Ramakrishna Kakarala. The bispectrum as a source of phase-sensitive invariants for fourier descrip-  
550 tors: a group-theoretic approach, 2012. URL <https://arxiv.org/abs/0902.0196>.  
551
- 552 Risi Kondor. A novel set of rotationally and translationally invariant features for images based on  
553 the non-commutative bispectrum, 2007. URL <https://arxiv.org/abs/cs/0701127>.  
554
- 555 S.G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE*  
556 *Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989. doi: 10.1109/  
557 34.192463.
- 558 Stéphane Mallat. Group invariant scattering. *Communications on Pure and Applied Math-*  
559 *ematics*, 65(10):1331–1398, 2012. doi: <https://doi.org/10.1002/cpa.21413>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.21413>.  
560
- 561 Stéphane Mallat, Sixin Zhang, and Gaspar Rochette. Phase harmonic correlations and convolu-  
562 tional neural networks. *Information and Inference: A Journal of the IMA*, 9(3):721–747, 11  
563 2019. ISSN 2049-8772. doi: 10.1093/imaiai/iaz019. URL [https://doi.org/10.1093/](https://doi.org/10.1093/imaiai/iaz019)  
564 [imaiai/iaz019](https://doi.org/10.1093/imaiai/iaz019).  
565
- 566 Dylan Nelson, Volker Springel, Annalisa Pillepich, Vicente Rodriguez-Gomez, Paul Torrey, Shy  
567 Genel, Mark Vogelsberger, Ruediger Pakmor, Federico Marinacci, Rainer Weinberger, Luke  
568 Kelley, Mark Lovell, Benedikt Diemer, and Lars Hernquist. The IllustrisTNG simulations:  
569 public data release. *Computational Astrophysics and Cosmology*, 6(1):2, May 2019. doi:  
570 10.1186/s40668-019-0028-x.
- 571 Julian Newman, Aleksandra Pidde, and Aneta Stefanovska. Defining the wavelet bispectrum.  
572 *Applied and Computational Harmonic Analysis*, 51:171–224, 2021. ISSN 1063-5203. doi:  
573 <https://doi.org/10.1016/j.acha.2020.10.005>. URL [https://www.sciencedirect.com/](https://www.sciencedirect.com/science/article/pii/S1063520320300750)  
574 [science/article/pii/S1063520320300750](https://www.sciencedirect.com/science/article/pii/S1063520320300750).  
575
- 576 C.L. Nikias and M.R. Raghuveer. Bispectrum estimation: A digital signal processing framework.  
577 *Proceedings of the IEEE*, 75(7):869–891, 1987. doi: 10.1109/PROC.1987.13824.
- 578 A.V. Oppenheim and J.S. Lim. The importance of phase in signals. *Proceedings of the IEEE*, 69(5):  
579 529–541, 1981. doi: 10.1109/PROC.1981.12022.
- 580
- 581 Edouard Oyallon, Stéphane Mallat, and Laurent Sifre. Generic deep networks with wavelet scat-  
582 tering. In Yoshua Bengio and Yann LeCun (eds.), *2nd International Conference on Learning*  
583 *Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Workshop Track Proceed-*  
584 *ings*, 2014. URL <http://arxiv.org/abs/1312.5940>.  
585
- 586 Badri Patro and Vijay Agneeswaran. Scattering vision transformer: Spectral mixing matters. In  
587 A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine (eds.), *Advances in*  
588 *Neural Information Processing Systems*, volume 36, pp. 54152–54166. Curran Associates, Inc.,  
589 2023. URL [https://proceedings.neurips.cc/paper\\_files/paper/2023/](https://proceedings.neurips.cc/paper_files/paper/2023/file/a97f8072e51a785434b2da3e9cbf5aae-Paper-Conference.pdf)  
590 [file/a97f8072e51a785434b2da3e9cbf5aae-Paper-Conference.pdf](https://proceedings.neurips.cc/paper_files/paper/2023/file/a97f8072e51a785434b2da3e9cbf5aae-Paper-Conference.pdf).
- 591 M. M. Pawlik, V. Wild, C. J. Walcher, P. H. Johansson, C. Villforth, K. Rowlands, J. Mendez-Abreu,  
592 and T. Hewlett. Shape asymmetry: a morphological indicator for automatic detection of galaxies  
593 in the post-coalescence merger stages. *Monthly Notices of the Royal Astronomical Society*, 456  
(3):3032–3052, March 2016. doi: 10.1093/mnras/stv2878.

- 594 Prasanna Porwal, Samiksha Pachade, Manesh Kokare, Girish Deshmukh, Jaemin Son, Woong Bae,  
595 Lihong Liu, Jianzong Wang, Xinhui Liu, Liangxin Gao, TianBo Wu, Jing Xiao, Fengyan Wang,  
596 Baocai Yin, Yunzhi Wang, Gopichandh Danala, Linsheng He, Yoon Ho Choi, Yeong Chan Lee,  
597 Sang-Hyuk Jung, Zhongyu Li, Xiaodan Sui, Junyan Wu, Xiaolong Li, Ting Zhou, Janos Toth,  
598 Agnes Baran, Avinash Kori, Sai Saketh Chennamsetty, Mohammed Safwan, Varghese Alex,  
599 Xingzheng Lyu, Li Cheng, Qin hao Chu, Pengcheng Li, Xin Ji, Sanyuan Zhang, Yaxin Shen, Ling  
600 Dai, Oindrila Saha, Rachana Sathish, Tânia Melo, Teresa Araújo, Balazs Harangi, Bin Sheng,  
601 Ruogu Fang, Debdoot Sheet, Andras Hajdu, Yuanjie Zheng, Ana Maria Mendonça, Shaoting  
602 Zhang, Aurélio Campilho, Bin Zheng, Dinggang Shen, Luca Giancardo, Gwenolé Quéllec, and  
603 Fabrice Mériaudeau. Idrid: Diabetic retinopathy – segmentation and grading challenge. *Med-  
604 ical Image Analysis*, 59:101561, 2020. ISSN 1361-8415. doi: [https://doi.org/10.1016/j.media.  
605 2019.101561](https://doi.org/10.1016/j.media.2019.101561). URL [https://www.sciencedirect.com/science/article/pii/  
606 S1361841519301033](https://www.sciencedirect.com/science/article/pii/S1361841519301033).
- 607 Rosemberg Rodriguez, Eva Dokladalova, and Petr Dokládál. Rotation invariant cnn using scattering  
608 transform for image classification. In *2019 IEEE International Conference on Image Processing  
609 (ICIP)*, pp. 654–658. IEEE, 2019.
- 610 Sophia Sanborn, Christian A Shewmake, Bruno Olshausen, and Christopher J. Hillar. Bispectral  
611 neural networks. In *The Eleventh International Conference on Learning Representations, 2023*.  
612 URL <https://openreview.net/forum?id=xnsg4pfKb7>.
- 613 Andrew K Saydjari and Douglas P Finkbeiner. Equivariant wavelets: Fast rotation and translation  
614 invariant wavelet scattering transforms. *IEEE Transactions on Pattern Analysis and Machine  
615 Intelligence*, 45(2):1716–1731, 2022.
- 617 Laurent Sifre and Stéphane Mallat. Rotation, scaling and deformation invariant scattering for texture  
618 discrimination. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1233–  
619 1240, 2013. doi: 10.1109/CVPR.2013.163.
- 620 Laurent Sifre and Stéphane Mallat. Rigid-motion scattering for texture classification, 2014. URL  
621 <https://arxiv.org/abs/1403.1687>.
- 623 E.P. Simoncelli, W.T. Freeman, E.H. Adelson, and D.J. Heeger. Shiftable multiscale transforms.  
624 *IEEE Transactions on Information Theory*, 38(2):587–607, 1992. doi: 10.1109/18.119725.
- 625 Nikolay Skarbnik, Yehoshua Y. Zeevi, and Chen Sagiv. The importance of phase in image pro-  
626 cessing. Technical report, Technion – Israel Institute of Technology, 2010. URL [https:  
627 //webee.technion.ac.il/people/zeevi/reports/773A.pdf](https://webee.technion.ac.il/people/zeevi/reports/773A.pdf).
- 629 B. Ph. van Milligen, E. Sánchez, T. Estrada, C. Hidalgo, B. Brañas, B. Carreras, and L. García.  
630 Wavelet bicoherence: A new turbulence analysis tool. *Physics of Plasmas*, 2(8):3017–3032, 08  
631 1995. ISSN 1070-664X. doi: 10.1063/1.871199. URL [https://doi.org/10.1063/1.  
632 871199](https://doi.org/10.1063/1.871199).
- 633 Irene Waldspurger, Alexandre d’Aspremont, and Stéphane Mallat. Phase recovery, maxcut and  
634 complex semidefinite programming. *Mathematical Programming*, 149(1):47–81, 2015.
- 636 Mike Walmsley, Micah Bowles, Anna M. M. Scaife, Jason Shingirai Makechemu, Alexander J.  
637 Gordon, Annette M. N. Ferguson, Robert G. Mann, James Pearson, Jürgen J. Popp, Jo Bovy, Josh  
638 Speagle, Hugh Dickinson, Lucy Fortson, Tobias Géron, Sandor Kruk, Chris J. Lintott, Kameswara  
639 Mantha, Devina Mohan, David O’Ryan, and Inigo V. Slijepevic. Scaling Laws for Galaxy Images.  
640 *arXiv e-prints*, art. arXiv:2404.02973, April 2024. doi: 10.48550/arXiv.2404.02973.
- 641 Mengwan Wei, Yongzhao Du, Xiuming Wu, Qichen Su, Jianqing Zhu, Lixin Zheng, Guorong Lv,  
642 and Jiafu Zhuang. A benign and malignant breast tumor classification method via efficiently  
643 combining texture and morphological features on ultrasound images. *Computational and mathe-  
644 matical methods in medicine*, 2020(1):5894010, 2020.
- 646 Daniel E. Worrall, Stephan J. Garbin, Daniyar Turmukhambetov, and Gabriel J. Brostow. Harmonic  
647 networks: Deep translation and rotation equivariance. In *Proceedings of the IEEE Conference on  
Computer Vision and Pattern Recognition (CVPR)*, July 2017.

648 Jiancheng Yang, Rui Shi, and Bingbing Ni. Medmnist classification decathlon: A lightweight automl  
649 benchmark for medical image analysis. In *IEEE 18th International Symposium on Biomedical*  
650 *Imaging (ISBI)*, pp. 191–195, 2021.  
651

652 Jiancheng Yang, Rui Shi, Donglai Wei, Zequan Liu, Lin Zhao, Bilian Ke, Hanspeter Pfister, and  
653 Bingbing Ni. Medmnist v2-a large-scale lightweight benchmark for 2d and 3d biomedical image  
654 classification. *Scientific Data*, 10(1):41, 2023.

655 John Zarka, Louis Thiry, Tomás Angles, and Stéphane Mallat. Deep network classification by scat-  
656 tering and homotopy dictionary learning, 2020. URL [https://arxiv.org/abs/1910.](https://arxiv.org/abs/1910.03561)  
657 [03561](https://arxiv.org/abs/1910.03561).

658 S. Kevin Zhou, Hayit Greenspan, Christos Davatzikos, James S. Duncan, Bram Van Ginneken,  
659 Anant Madabhushi, Jerry L. Prince, Daniel Rueckert, and Ronald M. Summers. A review of  
660 deep learning in medical imaging: Imaging traits, technology trends, case studies with progress  
661 highlights, and future promises. *Proceedings of the IEEE*, 109(5):820–838, 2021. doi: 10.1109/  
662 JPROC.2021.3054390.  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701

## A TECHNICAL APPENDICES AND SUPPLEMENTARY MATERIAL

### A.1 DATASETS

We plot visualisations of MedMNIST samples in figure 3. In each task, relevant occur at varying scales. For example, BreastMNIST relies on broader symmetries such as regularity of the tumour.

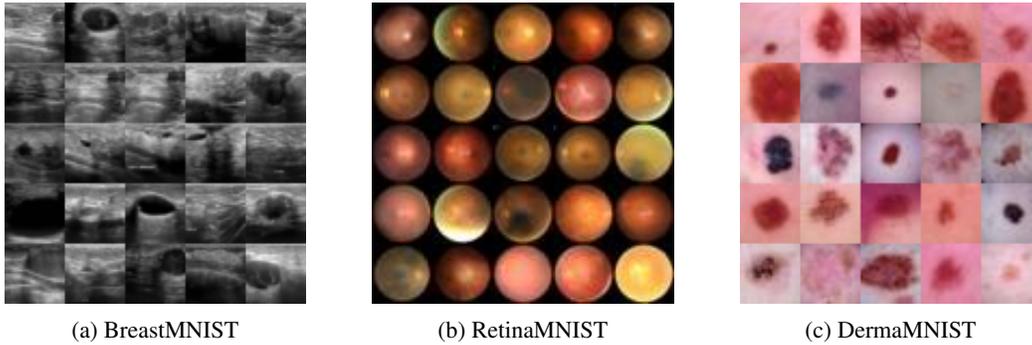


Figure 3: MedMNIST dataset samples

In figure 4 we present a selection of random samples from our galaxy dataset. Evidently, there is no observational noise. Similarly, one can observe the presence of merger features; for example, the top-left sample depicts tidal tails.

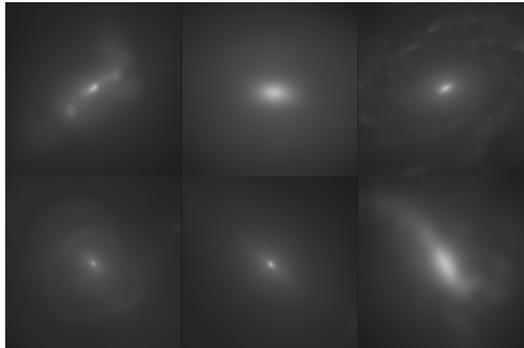


Figure 4: Galaxy merger samples.

Finally, samples from KTH-TIPS (figure 5) are comparatively simpler. Classes are distinct, and discrimination is subject to solely textural patterns at varying scales.

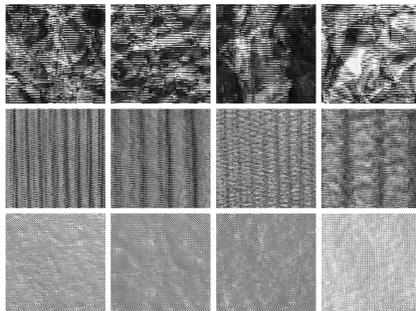
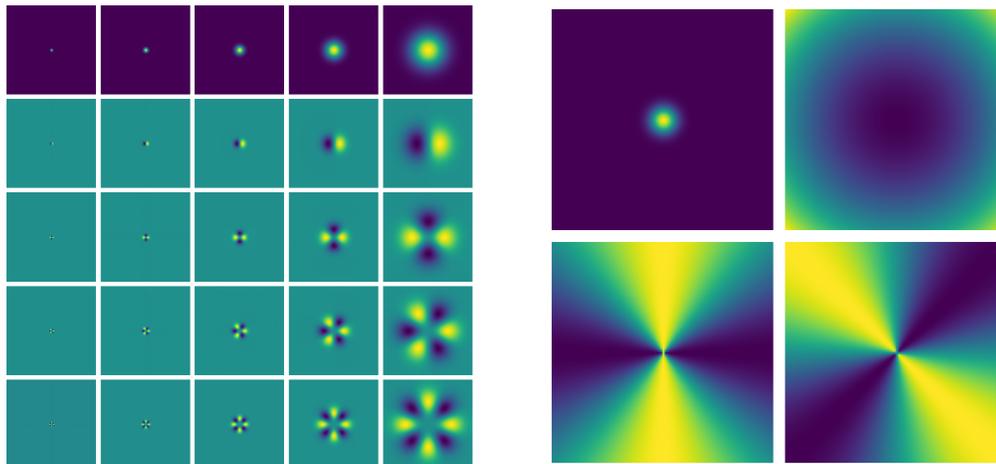


Figure 5: Samples from KTH-TIPS. Each row is a different class.

## A.2 OPERATORS

Figure 6a shows solid harmonic wavelets over varying scales (horizontally) and angular frequencies (vertically). Figure 6b shows the components that build up solid harmonic wavelets from Eq. 22. The Gaussian window,  $\frac{1}{\sqrt{(2\pi\sigma^2)}}e^{-r^2/2\sigma^2}$  plotted on the top left is used to localise the wavelet support, while the radial component,  $r^\ell$  top right, removes a singularity at 0. Finally, the complex frequency components,  $e^{i\ell\varphi}$  with real and imaginary plotted at bottom left and right respectively, span the angular coordinate of the wavelet.

$$\psi_{\sigma,\ell}(r,\varphi) = \frac{1}{\sqrt{(2\pi\sigma^2)}}e^{-\frac{r^2}{2\sigma^2}}r^\ell e^{i\ell\varphi} \quad (22)$$



(a) Real parts of solid harmonic filter bank. J increases left-to-right. L increases top-to-bottom.

(b) Decomposition of wavelet. Top left to bottom right: Gaussian window, radial, real and imaginary angular components.

Figure 6: Solid harmonic wavelets in 2D

Figure 7 compares wavelet responses for both Morlet (Andreux et al., 2018) and Solid Harmonic filters. While Morlet wavelets effectively extract prominent features, such as edges, solid harmonic wavelets excel in encoding textural information – in addition to edges.

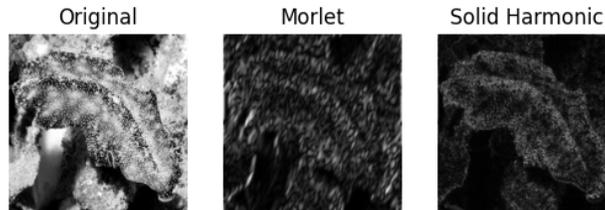
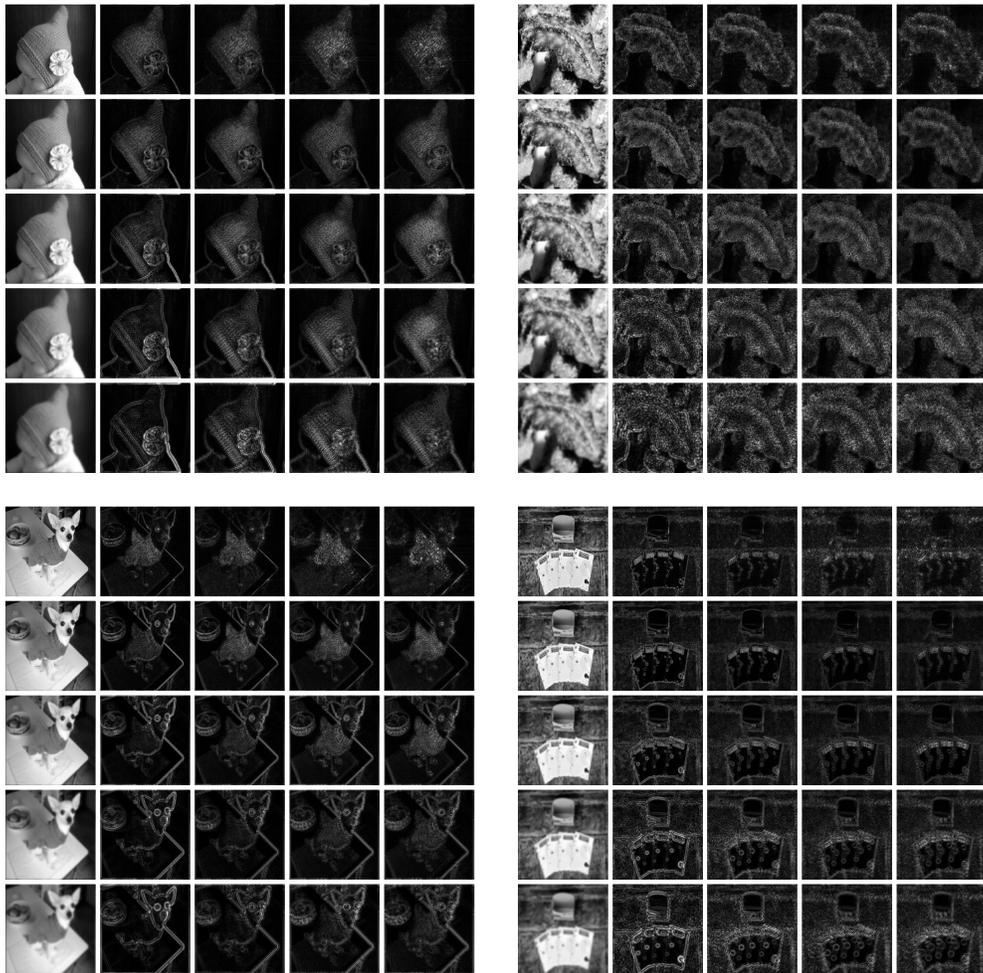


Figure 7: Wavelet transform responses using Morlet and Solid Harmonic wavelets

We plot how extracted features from the solid harmonic wavelet transform vary with solid harmonic wavelet parametrisation in figure 8. By increasing  $\ell$ , the features extracted by the solid harmonic wavelet transform transition from edges to texture, such that using dilated wavelets one could encode both textures and edges across scales. Thus, solid harmonic wavelet responses more effectively decouple these feature types. One can reason that since increasing the angular frequency produces more sinusoidal oscillations, solid harmonic wavelet transforms with increasing  $\ell$  can capture finer-scale variations in texture as a result of increased sensitivity to directional variations in the input

810 signal. Simultaneously, increasing the scale  $j$  captures coarser, higher-level features. To this extent,  
 811 by building a filter bank of dilated solid harmonic wavelets one can produce a complete represen-  
 812 tation which encompasses edges and textures of varying granularity. Contrasting to Morlet wavelet  
 813 transforms in figure 7, we demonstrate that solid harmonic wavelet transforms capture an increas-  
 814 ingly diverse collection of feature representations.



853 Figure 8: Effect of  $\ell$  and  $j$  parameters on features produced by solid harmonic wavelet transform.  
 854 Left-to-right:  $\ell$  increases from 0 (Gaussian) to 10 in steps of 2. Top-to-bottom:  $j$  increases.

855  
856  
857  
858  
859  
860  
861  
862  
863 In figure 9 we plot the magnitude for  $L_1$  pooling of SHWBic coefficients. One can observe similarity  
 between samples belonging to the same class (babies).

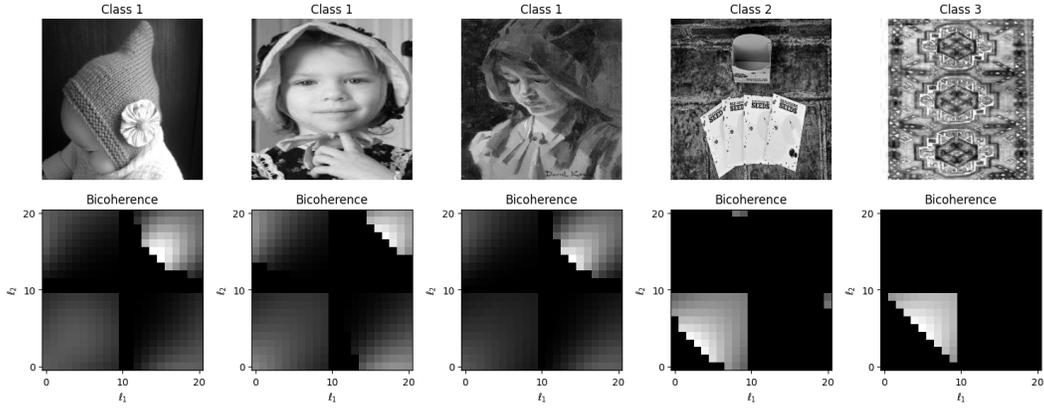


Figure 9: Plots for SHWBic  $L_1$  pooling magnitude for different ImageNet classes. Note intra-class similarity.

### A.3 BISPECTRUM COVARIANCE

This appendix establishes that SHWB and SHWSB are covariant to roto-translations under the special Euclidean group  $SE(2)$ , and that integration over spatial frequencies yields invariants. To establish the covariance of the solid harmonic wavelet bispectrum (SHWB) under rotations in Eq. 23, we begin by considering the rotational behavior of the solid harmonic wavelet transform. For a planar rotation  $R_\gamma$ , the wavelet coefficients transform as shown in Eq. 8:

$$W_{j,\ell}(R_\gamma x) = e^{i\ell\gamma} R_\gamma W_{j,\ell}(x),$$

where  $R_\gamma$  denotes the spatial rotation operator acting on the coordinates, and the phase term  $e^{i\ell\gamma}$  arises from the equivariance of solid harmonics under rotation of the input signal.

We substitute Equation 8 into the definition of the SHWB:

$$\text{SHWB}[x, j, \ell_1, \ell_2] = W_{j,\ell_1}(x) \cdot W_{j,\ell_2}(x) \cdot \overline{W_{j,\ell_1+\ell_2}(x)},$$

Applying a rotation  $R_\gamma$  to the input  $x$  and using Equation 8, we have:

$$\begin{aligned} \text{SHWB}[R_\gamma x, j, \ell_1, \ell_2] &= W_{j,\ell_1}(R_\gamma x) \cdot W_{j,\ell_2}(R_\gamma x) \cdot \overline{W_{j,\ell_1+\ell_2}(R_\gamma x)} \\ &= e^{i\ell_1\gamma} R_\gamma W_{j,\ell_1}(x) \cdot e^{i\ell_2\gamma} R_\gamma W_{j,\ell_2}(x) \cdot \overline{e^{i(\ell_1+\ell_2)\gamma} R_\gamma W_{j,\ell_1+\ell_2}(x)} \\ &= e^{i(\ell_1+\ell_2-(\ell_1+\ell_2))\gamma} R_\gamma \left( W_{j,\ell_1}(x) \cdot W_{j,\ell_2}(x) \cdot \overline{W_{j,\ell_1+\ell_2}(x)} \right) \\ &= R_\gamma \text{SHWB}[x, j, \ell_1, \ell_2]. \end{aligned} \quad (23)$$

Thus, the SHWB is *covariant* to rotations. Crucially, this covariance hinges on the angular momentum conservation condition  $\ell_3 = \ell_1 + \ell_2$ , which ensures that the global phase factor  $e^{i(\ell_1+\ell_2-\ell_3)\gamma}$  cancels. This cancellation is what distinguishes the SHWB from classical wavelet scattering: the phase information is preserved without requiring a modulus to enforce rotational invariance.

Since the underlying wavelet convolutions are also covariant to translations Mallat (2012), the SHWB inherits translation covariance. Full *roto-translation invariance* is then achieved by applying a suitable *reduction operator*, such as spatial averaging over the domain.

This formalism justifies the invariance properties of SHWB and illustrates how higher-order interactions can retain geometric structure without sacrificing phase information, offering a more expressive alternative to classical scattering transforms.

Finally, we note that the solid harmonic wavelets  $\psi_{\ell,j}$ , producing the responses  $W_{j,\ell}(x)$ , are orthogonal across different scales  $j$  and angular frequencies  $\ell$ . As a result, the SHWB generates a set

918 of largely independent invariants. This orthogonality ensures that while irrelevant variations (e.g.,  
919 due to transformations) are removed, the essential discriminative information in the signal is largely  
920 preserved. Consequently, the SHWB forms a rich invariant representation that captures higher-order  
921 geometric structure without discarding informative phase content.  
922

923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971