

Automated Spiral-Arm Pitch-Angle Estimation in Galaxy Images Using Deep Learning

Ramil R. Yavaev

Volgograd State University

Maxim V. Polyakov

m.v.polyakov@volsu.ru

Volgograd State University

Research Article

Keywords: spiral galaxies, galaxy morphology, spiral arms, pitch angle, deep learning, convolutional neural networks, synthetic data, Sloan Digital Sky Survey

Posted Date: October 13th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11058145/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Automated Spiral-Arm Pitch-Angle Estimation in Galaxy Images Using Deep Learning

Ramil R. Yavaev¹ and Maxim V. Polyakov^{1*}

^{1*}Department of Information Systems and Computer Modelling,
Institute of Mathematics and Information Technology, Volgograd State
University, Volgograd, Russia.

*Corresponding author(s). E-mail(s): m.v.polyakov@volsu.ru;

Abstract

Spiral-arm pitch angle is a compact quantitative descriptor of spiral-galaxy structure and is related to galaxy morphology, kinematics, and central mass concentration. Its large-scale use is nevertheless limited by the scarcity of homogeneous numerical labels and by the sensitivity of conventional measurements to deprojection, noise, dust, and arm fragmentation. We present an end-to-end deep-learning pipeline for continuous pitch-angle estimation from galaxy images. A parametric three-dimensional generator produces spiral galaxies with exact pitch-angle labels and controlled variations in arm number, viewing inclination, dust content, and flocculence. Images are processed by automatic disc deprojection followed by spatial-frequency enhancement using unsharp masking, contrast-limited adaptive histogram equalisation, and Fourier-domain filtering. A MobileNetV2 backbone pretrained on ImageNet is then adapted in two stages for regression. On 2,200 synthetic test images spanning pitch angles of 5° – 35° , the full pipeline achieved a mean absolute error of 1.01° and a root mean square error of 1.34° . External validation on 65 Sloan Digital Sky Survey galaxies with literature pitch-angle measurements reduced the mean absolute error from 7.07° without the full preprocessing pipeline to 5.83° , while the root mean square error decreased from 10.42° to 7.61° . The results show that geometry-aware preprocessing substantially improves synthetic-to-real transfer, although residual compression of the predicted pitch-angle range indicates that a larger real labelled sample and explicit treatment of bars and non-logarithmic arms remain necessary.

Keywords: spiral galaxies, galaxy morphology, spiral arms, pitch angle, deep learning, convolutional neural networks, synthetic data, Sloan Digital Sky Survey

1 Introduction

Galaxy morphology has provided a compact observational language for describing galaxy structure since the classical Hubble sequence (Hubble, 1926). In disc galaxies, the presence and organisation of spiral arms remain among the most conspicuous morphological features, but their physical origin and long-term evolution are still debated (Dobbs & Baba, 2014). Modern visual catalogues record not only broad Hubble type but also bars, rings, arm multiplicity, flocculence, and other structural details (Nair & Abraham, 2010; Willett et al., 2013). Importantly, these properties are not interchangeable. In particular, the winding of the spiral pattern contains information that is partly independent of bulge prominence and of the traditional Hubble type; large modern samples show substantial scatter between arm winding and bulge-related morphology (Masters et al., 2019). This motivates treating spiral-arm pitch angle as a quantitative structural parameter in its own right rather than merely as a visual subclassification criterion.

The spiral-arm pitch angle is defined as the angle between the tangent to a spiral arm and the tangent to a circle centred on the galaxy nucleus. Small pitch angles correspond to tightly wound spirals, whereas large pitch angles describe more open patterns. For a logarithmic spiral the pitch angle is constant with radius, which makes it a convenient scalar descriptor of arm geometry. Observational studies have linked pitch angle to several physical properties of galaxies, including Hubble type, bulge prominence, stellar mass, central velocity dispersion, rotation-curve morphology, and the central concentration of mass (Seigar & James, 1998; Seigar et al., 2006; Davis et al., 2015; Yu & Ho, 2019). Numerical experiments likewise indicate that the winding of spiral structure can respond to galactic shear and the underlying mass distribution (Grand et al., 2013). The proposed connection between pitch angle and central mass concentration has also motivated the use of spiral morphology as an imaging-based proxy for supermassive black-hole mass (Seigar et al., 2008; Berrier et al., 2013; Davis et al., 2017). These correlations are not universally recovered: larger or differently selected samples have reported weaker relations with shear, bulge mass, or bar strength (Díaz-García et al., 2019; Yu & Ho, 2019). Consequently, homogeneous and reproducible pitch-angle measurements over large samples are important not only for catalogue construction but also for testing which reported scaling relations are robust to sample selection and measurement methodology.

Pitch angle is also directly relevant to the longstanding problem of spiral-arm formation. Classical density-wave theory interprets spiral structure as a long-lived wave pattern propagating through the stellar disc (Lin & Shu, 1964), whereas numerical work has demonstrated transient, recurrent, tidally driven, and self-perpetuating alternatives (D’Onghia et al., 2013; Dobbs & Baba, 2014). Self-consistent collisionless-gaseous simulations of a multi-component Milky Way-like disc have further shown that a central bar can coexist with a persistent multi-armed spiral pattern represented by a superposition of two-, three-, and four-armed modes over several gigayears (Khrapov et al., 2021). In such scenarios, the instantaneous pitch angle and its radial or temporal evolution can provide constraints on arm dynamics. Simulations have shown a relation between pitch angle and shear (Grand et al., 2013), while observational analyses have demonstrated that real galaxies often depart from the idealised picture of a

single logarithmic spiral. Pitch angles can differ between arm segments and can vary substantially within the same galaxy (Savchenko & Reshetnikov, 2013; Honig & Reid, 2015; Díaz-García et al., 2019; Lingard et al., 2021). A single global pitch angle should therefore be interpreted as an effective summary of the visible spiral pattern. Despite this simplification, such a descriptor remains useful for population-level studies, initial catalogue characterisation, and selection of galaxies for subsequent arm-by-arm analysis.

The need for automated measurements is becoming increasingly acute as imaging surveys continue to expand in area, depth, and cadence. The Sloan Digital Sky Survey (SDSS) established the scale at which homogeneous digital imaging could transform galaxy-population studies (York et al., 2000; Abazajian et al., 2009); the Vera C. Rubin Observatory Legacy Survey of Space and Time is expected to increase the volume of optical imaging by orders of magnitude (Ivezić et al., 2019). Large visual classification projects partly addressed the resulting annotation bottleneck. Galaxy Zoo and Galaxy Zoo 2 demonstrated that citizen scientists can provide reliable morphological information for hundreds of thousands of galaxies (Lintott et al., 2008; Willett et al., 2013), while later projects have combined volunteer labels with deep learning to extend detailed morphology measurements to still larger imaging samples (Walmsley et al., 2022). Nevertheless, these catalogues primarily provide categorical or ordinal responses—for example, whether spiral structure is present, how tightly wound the arms appear, or how many arms are visible. Such labels are extremely valuable for morphology studies but do not provide a homogeneous continuous pitch angle in degrees for every object.

Several specialised techniques have therefore been developed for quantitative pitch-angle measurement. Two-dimensional Fourier decomposition identifies dominant logarithmic spiral modes and has been extensively used to estimate global pitch angle (Davis et al., 2012); related Fourier approaches can also examine radial variation in spiral geometry (Savchenko & Reshetnikov, 2013). SpArcFiRe automatically identifies arm segments and fits logarithmic spiral arcs to them (Davis & Hayes, 2014), and its measurements have been combined with Galaxy Zoo information to analyse arm number and winding in thousands of SDSS spirals (Hart et al., 2017). Template-based methods such as Spirality compare galaxy images with spiral patterns of known pitch (Shields et al., 2022), whereas citizen-science arm traces have been incorporated into probabilistic pitch-angle inference (Lingard et al., 2021). Comparisons among specialised algorithms show that the inferred angle can depend on deprojection, radial range, segmentation, arm irregularity, and other methodological choices (Hewitt & Treuthardt, 2020). Thus, even though classical methods are physically interpretable, obtaining consistent measurements for large heterogeneous image collections remains non-trivial.

Deep learning provides an alternative route because convolutional neural networks (CNNs) learn morphological features directly from image pixels. Rotation-aware CNNs achieved strong performance on Galaxy Zoo morphology questions (Dieleman et al., 2015), and subsequent studies demonstrated that deep networks can reproduce detailed visual classifications across large SDSS and Hubble samples

(Huertas-Company et al., 2015; Domínguez Sánchez et al., 2018). Comparative studies have further established the competitiveness of CNN-based approaches relative to conventional machine-learning descriptors for galaxy morphology (Barchi et al., 2020), while modern Galaxy Zoo releases use deep learning to propagate volunteer classifications to hundreds of thousands of galaxies (Walmsley et al., 2022). More specialised recent models can identify spiral-arm multiplicity directly from images (Lee et al., 2025). However, direct end-to-end regression of a continuous spiral-arm pitch angle remains much less common than broad morphological classification. The difficulty is not primarily the regression formulation itself, but the scarcity of large, homogeneous samples with numerical pitch-angle labels.

Synthetic data offer a practical means of addressing this label scarcity because the desired geometric parameter can be fixed exactly during image generation. Synthetic or generative astronomical images have increasingly been investigated for survey simulation and machine-learning development (Smith & Geach, 2019). At the same time, transferring a model from one image domain to another is a well-known machine-learning problem: a network trained on a source distribution can learn domain-specific cues that do not generalise to the target distribution (Pan & Yang, 2010). This issue is particularly relevant in astronomy, where changes in survey depth, point-spread function, noise, background structure, and photometric properties can substantially alter the appearance of the same morphological class. Cross-survey studies have therefore begun to apply explicit domain-adaptation methods to galaxy morphology (Ćiprijanović et al., 2023). Data augmentation can improve robustness under limited training data (Shorten & Khoshgoftaar, 2019), but it cannot by itself guarantee that a synthetic generator reproduces the full statistics of real telescope images. For pitch-angle regression, projection effects create an additional domain mismatch because an inclined disc changes the apparent geometry of the spiral pattern in a predictable but potentially large way.

The present study addresses this synthetic-to-real problem by combining a controllable synthetic galaxy generator, geometry-aware preprocessing, and transfer learning. An earlier proof-of-concept configuration used a frozen MobileNet backbone trained on synthetic images and showed that direct pitch-angle regression was feasible, reaching a synthetic test mean absolute error (MAE) of 1.52° ; however, the MAE increased to 7.94° on a set of 65 real SDSS galaxies with literature measurements. This discrepancy motivated a revised pipeline specifically designed to improve transfer to real telescope images. The final method uses a richer three-dimensional generator, a two-pass rendering strategy for exact labels under arbitrary inclinations, automatic disc deprojection, spatial-frequency contrast enhancement, and two-stage fine-tuning of MobileNetV2.

The main contributions of this work are therefore: (i) a parameterised synthetic-data generator that produces exact continuous pitch-angle labels while varying nuisance morphology and viewing geometry; (ii) an automatic preprocessing pipeline that explicitly compensates for disc inclination and enhances low-contrast arm structure; (iii) an ImageNet-pretrained MobileNetV2 regressor adapted for continuous pitch-angle prediction; and (iv) independent validation on real SDSS galaxies with pitch angles reported by Yu & Ho (2019). The central question is whether such

physically and geometrically motivated preprocessing can reduce the synthetic-to-real performance gap without requiring a large real labelled training set.

The remainder of this paper is organised as follows. Section 2 describes the study design, data sources, synthetic galaxy generation, image preprocessing, and the MobileNetV2-based regression model. Section 3 presents the results of the computational experiments on synthetic and real SDSS galaxy images. Section 4 discusses the performance of the proposed approach, its limitations, and its potential applications. Finally, Section 5 summarises the main conclusions and outlines directions for future work.

2 Materials and methods

2.1 Study design and data sources

The study was organised around a synthetic training domain and an independent real validation domain (Fig. 1). Synthetic images were used because the pitch angle is known exactly from the generative model, allowing the regression target to be assigned without human measurement noise. Real images were used only for external evaluation of transfer to observational data.

The final synthetic dataset contained approximately 11,000 rendered galaxy images. It was divided into training, validation, and test subsets in a 70:10:20 proportion, giving 2,200 images in the held-out test subset. The target pitch angle ψ covered 5° to 35° , a range that contains the typical values reported for nearby spiral galaxies while providing examples of both tightly and loosely wound structures. The generator also varied other morphological factors so that the network could not solve the task by relying on a single fixed visual template.

External validation used 65 spiral galaxies from SDSS for which independent pitch-angle measurements were available from Yu & Ho (2019). That study analysed 79 galaxies in total; the present validation set comprised the 65 objects for which the image preparation and matching procedure used in this project yielded suitable inputs. The SDSS images were centred on the galaxy nucleus, cropped to 600×600 pixels, and stored without additional lossy compression before preprocessing. The literature values were not used during optimisation of the network weights.

2.2 Parametric generation of spiral galaxies

The synthetic generator was implemented in Python using a three-dimensional rendering workflow based on Ursina and PyOpenGL. The geometry of each spiral arm was represented by a logarithmic spiral,

$$r(\theta) = r_0 \exp[\theta \tan(\psi)], \quad (1)$$

where r is radial distance, θ is azimuthal angle, r_0 is a reference radius, and ψ is the pitch angle. The use of a logarithmic form follows the standard approximation adopted in many quantitative spiral-structure studies, while the limitations of assuming a single global pitch angle are considered in Sect. 4.

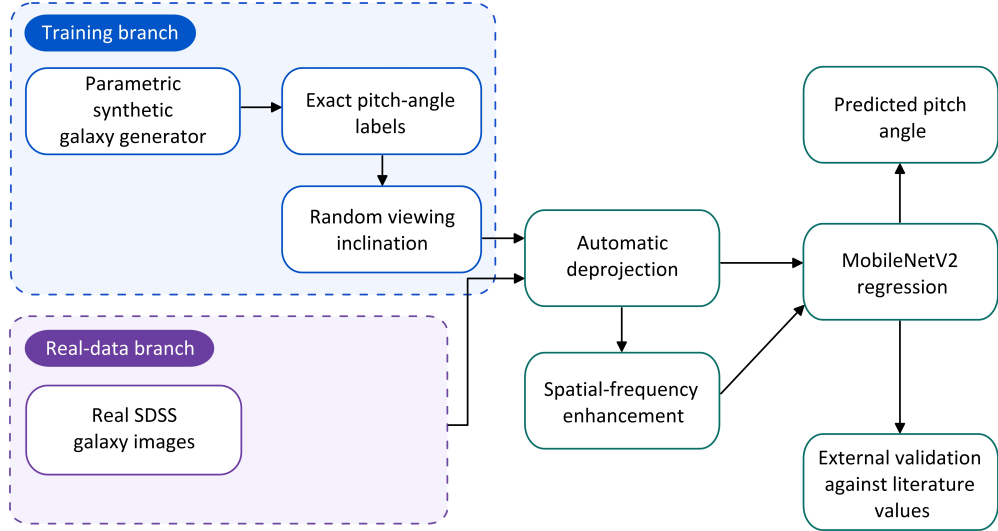


Fig. 1 Overview of the synthetic-to-real pitch-angle regression pipeline. Synthetic galaxies provide exact pitch-angle labels and are rendered under variable viewing inclinations. Synthetic and real SDSS images pass through the same automatic deprojection and spatial-frequency enhancement stages before MobileNetV2 inference. Real-data predictions are evaluated against independent literature measurements.

For each synthetic galaxy, the generator sampled the number of spiral arms from 1 to 8 and varied the pitch angle together with nuisance properties such as bulge appearance, dust content, flocculence, colour preset, and disc inclination. The inclination i was sampled between 0° and 90° . Approximately 15,000 particles were distributed along the modelled spiral trajectories to represent luminous stellar concentrations and star-forming regions. Radial changes in particle density and brightness were introduced to avoid uniform arms, while short asymmetric segments and stochastic displacements were used to emulate fragmented or flocculent patterns.

A two-pass rendering procedure was used to separate the target label from projection effects. First, the virtual camera rendered the galaxy face-on ($i = 0^\circ$), for which the underlying logarithmic geometry and the assigned pitch angle were unambiguous. The target value was stored at this stage. The same three-dimensional galaxy was then rendered at the sampled inclination to generate the actual training image. Thus, the network received projected images but retained the exact intrinsic pitch-angle label. This design was intended to expose the model to realistic orientation variability without corrupting the ground truth.

Figure 2 shows representative synthetic images and real SDSS images used during development. The generator does not attempt a full radiative-transfer simulation; instead, it aims to reproduce large-scale geometric invariants relevant to spiral-arm recognition. Photometric realism and geometric fidelity therefore represent distinct objectives. The generator was deliberately used as a source of controlled morphology rather than as a physical forward model of all galaxy observables.

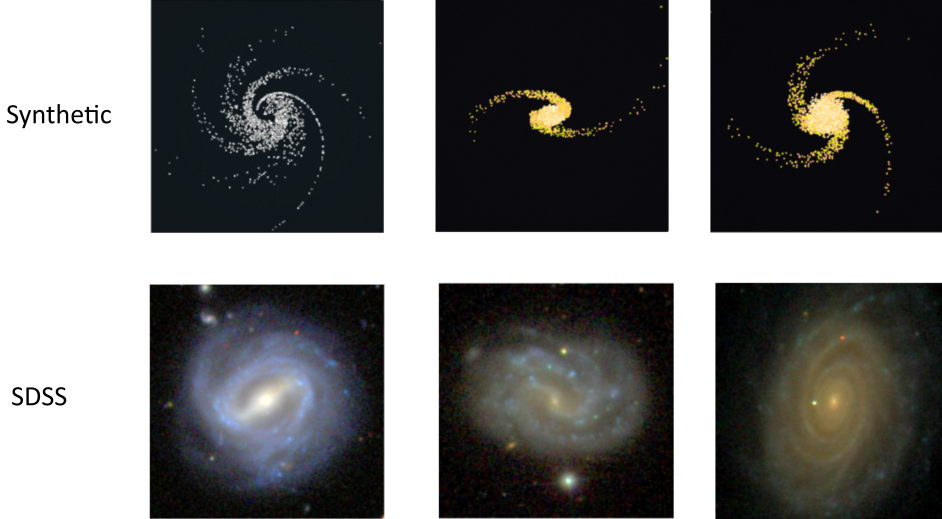


Fig. 2 Examples of synthetic spiral galaxies (top row) and real SDSS spiral galaxies (bottom row). The synthetic generator emphasises controllable large-scale arm geometry rather than full photometric realism.

2.3 Automatic geometric deprojection

Projection is a major source of bias in pitch-angle measurement because an intrinsically circular disc is observed as an ellipse when inclined relative to the line of sight. A CNN can in principle learn approximate invariance to this transformation, but doing so consumes model capacity and requires a sufficiently representative training set. We therefore incorporated explicit deprojection before feature extraction.

Each image was first converted to greyscale and segmented using Otsu thresholding. Morphological closing with a 5×5 kernel was applied to suppress small gaps and fragmented regions. An ellipse was then fitted to the largest detected contour, returning a centre (c_x, c_y) , major and minor axes (a, b) , and a position angle. Under the thin-disc approximation, the apparent axis ratio was used to estimate inclination,

$$i = \arccos\left(\frac{b}{a}\right), \quad s_y = \frac{a}{b}, \quad (2)$$

where s_y is the scale factor used to restore the compressed axis. A composite affine transformation translated the galaxy centre to the origin, rotated to align the ellipse, applied the anisotropic scaling, reversed the rotation, and translated the image back. The corrected image was cropped to a region extending approximately $1.5a$ around the fitted centre and resized to 224×224 pixels for MobileNetV2.

This image-based deprojection is deliberately simple and fully automatic. It does not incorporate a three-dimensional bulge model or intrinsic disc thickness. The advantage is scalability; the disadvantage is that barred, disturbed, or nearly edge-on systems can produce imperfect ellipse fits. These cases are discussed as a principal limitation of the present method.

2.4 Spatial–frequency enhancement of spiral structure

After deprojection, a second stage was used to improve visibility of arm structure under uneven background illumination and detector noise. Processing was performed primarily on the luminance channel in the CIELAB colour space so that local contrast could be modified without unnecessarily altering colour information. Unsharp masking first emphasised arm boundaries. Contrast-limited adaptive histogram equalisation (CLAHE) was then applied with a clip limit of 1.5 and an 8×8 local window. These values were chosen empirically during development: stronger equalisation created artificial breaks in synthetic arms, whereas weaker equalisation often failed to reveal low-contrast dust-lane and arm structure.

The CLAHE output was combined with a Fourier-domain component. A band-pass filtering step suppressed the very-low-frequency background and high-frequency detector noise while retaining intermediate spatial frequencies associated with the spiral pattern. The inverse transform was normalised and blended with the locally contrast-enhanced image. The final image was returned to RGB representation and scaled to the input range required by the neural network.

An example is shown in Fig. 3. Deprojection restores a more nearly face-on geometry, while the enhancement stage makes weak outer-arm segments more visible. Importantly, this operation is not intended to create morphology that is absent from the data; it is designed to increase the contrast of structures already present in the image.

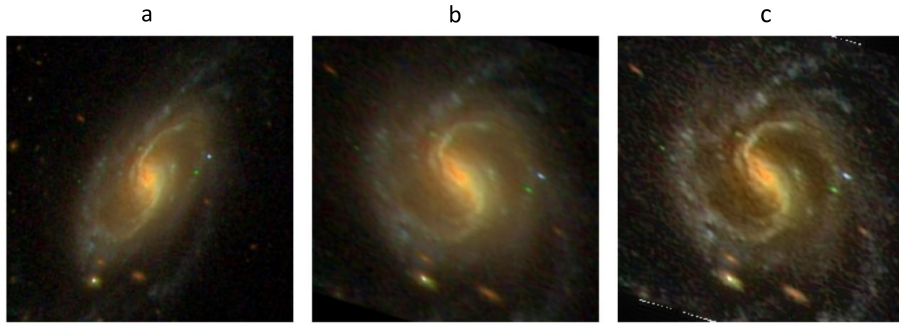


Fig. 3 Illustrative preprocessing sequence for a real spiral galaxy: (a) original image, (b) automatic geometric deprojection, and (c) spatial–frequency enhancement.

2.5 MobileNetV2 regression model and transfer learning

MobileNetV2 (Sandler et al., 2018) was selected as the convolutional backbone because depthwise separable convolutions provide a favourable compromise between representational capacity and computational cost. This property is useful for iterative experimentation and for potential future processing of large survey catalogues. The backbone was initialised with ImageNet weights (Deng et al., 2009) rather than trained from random initialisation.

The classification head of MobileNetV2 was removed and replaced by a regression head containing fully connected layers of 512, 256, and 128 units with rectified linear activation. Batch normalisation and dropout regularisation were used between dense stages. The output layer contained one neuron with linear activation, returning a continuous estimate of ψ in degrees.

Training proceeded in two phases. During the first phase, the MobileNetV2 feature extractor was frozen and only the regression head was optimised for 20 epochs using Adam and mean squared error (MSE) as the loss. Once validation performance reached a plateau, the upper 20–30 layers of the backbone were unfrozen for fine-tuning at the task-specific level while preserving lower-level ImageNet features. This staged procedure reduced instability observed when the entire backbone was unfrozen at once.

Training augmentation comprised random horizontal and vertical flips and brightness and contrast perturbations of approximately $\pm 20\%$. Arbitrary-angle rotation was not retained in the final augmentation policy because the explicit deprojection stage already handled orientation effects and additional image interpolation introduced visible edge artefacts. No augmentation was applied during validation or testing. On an NVIDIA RTX 3050 GPU, one training epoch required approximately 45 s in the reported implementation.

2.6 Evaluation metrics and ablation design

Regression performance was quantified primarily by MAE and root mean square error (RMSE),

$$\text{MAE} = \frac{1}{n} \sum_{k=1}^n |\hat{\psi}_k - \psi_k|, \quad (3)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{k=1}^n (\hat{\psi}_k - \psi_k)^2}, \quad (4)$$

where ψ_k and $\hat{\psi}_k$ are the reference and predicted pitch angles. Mean absolute percentage error (MAPE) was additionally reported for the synthetic test set. MAE is particularly interpretable here because it is expressed directly in degrees, whereas RMSE assigns greater weight to large outliers (Willmott & Matsuura, 2005).

The effect of preprocessing was assessed by comparing otherwise matched configurations with and without the complete deprojection and enhancement pipeline. On synthetic images, the relevant comparison was MAE 1.23° without the full pipeline versus 1.01° with it. On real SDSS images, MAE and RMSE were compared between basic resizing/normalisation and the full preprocessing pipeline. The relationship between synthetic inference error and generator-controlled flocculence was evaluated using a rank correlation; the reported Spearman coefficient was $\rho = 0.31$ with $p < 0.01$.

A preliminary frozen-backbone implementation from the earlier development stage is also reported as a historical baseline because it motivated the present pipeline. It achieved MAE 1.52° on synthetic test images and 7.94° on the same 65-galaxy real

validation set. The final comparisons, however, focus on the controlled ablation within the revised pipeline.

3 Results

3.1 Pitch-angle regression on synthetic data

The final MobileNetV2 model reproduced the continuous pitch angle accurately on the held-out synthetic test set. Across 2,200 images spanning 5° – 35° , the full pipeline reached MAE 1.01° , RMSE 1.34° , and MAPE 4.2% (Table 1). These values indicate that, within the controlled synthetic domain, the model learned the intended relation between the global spiral geometry and the continuous target rather than merely separating broad winding classes.

Table 1 Pitch-angle regression on the synthetic test set.

Metric	Full pipeline	Unit
MAE	1.01	degrees
RMSE	1.34	degrees
MAPE	4.2	%

The preprocessing ablation showed that the improvement was not solely a consequence of the network architecture. When the final regressor was evaluated without the full deprojection and frequency-enhancement sequence, synthetic MAE was 1.23° . Adding the complete preprocessing pipeline reduced MAE by 0.22° , or approximately 17.9%. Figure 4 also shows the earlier proof-of-concept value of 1.52° for context. The progressive reduction in error is consistent with the interpretation that explicit geometric normalisation allows the network to allocate more capacity to the arm pattern itself.

Training behaviour also improved after staged fine-tuning. During the frozen-backbone phase, the recorded MSE decreased from 45.2 to 3.8 and MAE from approximately 5.9° to 1.8° . Fine-tuning the upper part of MobileNetV2 further reduced the final MSE to 1.79 and MAE to 1.01° . Attempts to unfreeze the entire backbone at once were less stable, supporting the use of gradual adaptation for this relatively specialised image domain.

3.2 Sensitivity to flocculent structure

The controlled parameters of the synthetic generator allowed error to be examined as a function of morphology. Among the tested generator variables, flocculence showed the clearest association with pitch-angle error, with $\rho = 0.31$ and $p < 0.01$. For strongly fragmented synthetic spirals, MAE increased locally to approximately 1.45° . This behaviour is expected because the target is defined by a single underlying logarithmic spiral, whereas the rendered visual pattern may contain stochastic breaks and

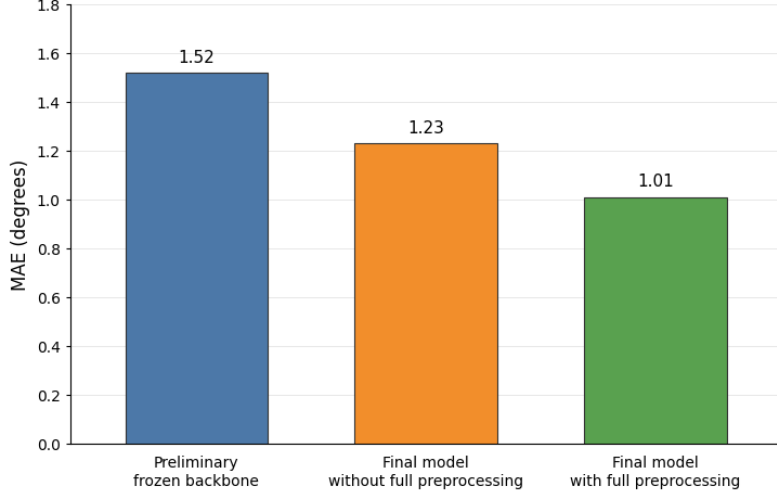


Fig. 4 Synthetic-domain MAE during development. The preliminary frozen-backbone experiment is shown for context; the controlled ablation of the final model is the comparison between the second and third bars.

asymmetric segments. The model must infer a global geometric trend from incomplete arm evidence.

Flocculent galaxies are precisely the systems for which a single global pitch angle can be physically and observationally ambiguous. Therefore, the increasing error should not be treated only as a machine-learning failure; it also marks a regime in which the target variable becomes less well represented by a single logarithmic spiral parameter.

3.3 External validation on real SDSS galaxies

The synthetic-domain accuracy did not transfer directly to real images, confirming a substantial domain gap. Nevertheless, the full preprocessing pipeline produced a consistent improvement on the independent sample of 65 SDSS galaxies with pitch angles from Yu & Ho (2019). Without the full pipeline, MAE was 7.07° and RMSE was 10.42° . After automatic deprojection and spatial-frequency enhancement, MAE decreased to 5.83° and RMSE to 7.61° (Table 2). The relative reduction was 17.5% for MAE and 27.0% for RMSE.

Table 2 External validation on 65 real SDSS spiral galaxies. The percentage improvement is calculated relative to the configuration without the full preprocessing pipeline.

Metric	Without full preprocessing	Full pipeline	Improvement
MAE (degrees)	7.07	5.83	17.5%
RMSE (degrees)	10.42	7.61	27.0%

The larger reduction in RMSE than in MAE indicates that preprocessing was particularly effective at reducing some of the larger deviations. Relative to the earlier frozen-backbone proof-of-concept (MAE 7.94°), the final pipeline reduced real-data MAE by approximately 26.6% (Fig. 5). This historical comparison combines several changes and should therefore not be interpreted as an isolated preprocessing effect; the controlled 17.5% ablation is the appropriate estimate for the contribution of the full preprocessing stage within the revised model.

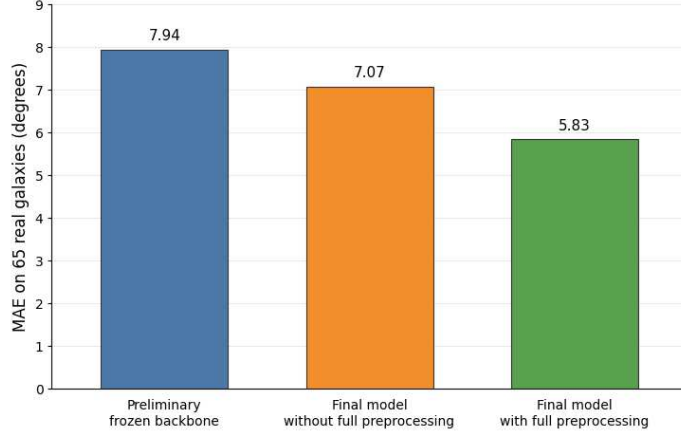


Fig. 5 MAE on the 65-galaxy real validation set during development. The preliminary value motivated the revised synthetic-to-real pipeline. The final controlled comparison is between the second and third bars.

Inspection of the original prediction plots showed a tendency for real-data predictions to occupy a narrower interval than the literature values, with low pitch angles being overestimated and high pitch angles being underestimated. This regression-to-the-mean behaviour is not reflected fully by a single MAE value and indicates that the remaining synthetic-to-real gap is scientifically important. The model is therefore better interpreted as a proof of concept for rapid continuous pitch-angle inference than as a drop-in replacement for established precision measurement techniques.

4 Discussion

The principal result of this study is that a CNN can be trained to predict a continuous spiral-arm pitch angle directly from galaxy images and can retain a measurable degree of accuracy when transferred from synthetic training data to real SDSS observations. This task differs from the more common deep-learning formulation of galaxy morphology as categorical classification. CNN-based classifiers can distinguish broad morphology or arm-number classes effectively (Dieleman et al., 2015; Lee et al., 2025), but a numerical pitch angle requires the network to preserve information about the geometry of the spiral pattern on a continuous scale.

Most established automated pitch-angle methods encode this geometry explicitly. Fourier-based methods identify dominant logarithmic modes (Davis et al., 2012; Savchenko & Reshetnikov, 2013); SpArcFiRe segments individual arms and performs logarithmic-arc fitting (Davis & Hayes, 2014); Spirality compares images with spiral templates of known pitch (Shields et al., 2022); and Galaxy Builder analyses incorporate citizen-science traces and probabilistic inference (Lingard et al., 2021). These approaches have strong interpretability and can return information beyond a single global prediction. The proposed approach should therefore be regarded as complementary to these specialised methods. Its potential advantage is computational simplicity at inference time after training: a preprocessed image can be mapped directly to a continuous estimate without per-object manual tracing or iterative template fitting.

Previous studies further show that a single global pitch angle should be interpreted with caution. Savchenko & Reshetnikov (2013) found substantial radial pitch-angle variation in many galaxies, while SpArcFiRe analyses similarly demonstrate segment-to-segment differences (Davis & Hayes, 2014). A CNN trained on one global reference value necessarily learns an effective aggregate of the visible structure. This can still be useful for large-sample screening and as an initial condition for more detailed fitting, but future versions should ideally predict a distribution, arm-specific values, or a radial pitch-angle profile rather than a single scalar.

The controlled ablation supports the central hypothesis of this work: preprocessing that reflects known image geometry can improve transfer from synthetic to real data. On the 65-galaxy sample, the full pipeline reduced MAE from 7.07° to 5.83° and RMSE from 10.42° to 7.61° . The improvement is physically plausible because inclination changes the apparent spiral geometry in a systematic way. Although data augmentation can expose a network to projected examples, explicit deprojection removes a nuisance transformation before the regression task.

The spatial-frequency stage addresses a different part of the domain gap. Synthetic arms are comparatively clean even when flocculence and dust are introduced, whereas real SDSS images contain background gradients, detector noise, foreground sources, patchy star formation, and complex extinction. CLAHE and frequency filtering increase the visibility of weak arm segments without forcing the CNN to learn all low-level contrast corrections from the synthetic domain. The larger relative improvement in RMSE suggests that this step may reduce catastrophic failures on some difficult images, although object-level analysis is needed to identify which morphological subtypes benefit most.

Nevertheless, a considerable performance gap remains: MAE increases from 1.01° on synthetic images to 5.83° on real galaxies. The compressed dynamic range of real predictions indicates that the synthetic generator still does not span the full observational distribution. More photorealistic simulations could help, but realism alone is not sufficient. A more effective strategy may combine synthetic pretraining with supervised or semi-supervised fine-tuning on a substantially larger set of real galaxies with homogeneous pitch-angle measurements.

A scalable pitch-angle estimator is valuable because pitch angle is more than a visual descriptor. Yu & Ho (2019) showed that pitch angle is related to galaxy morphology, stellar mass, central velocity dispersion, and rotation-curve shape. Davis et

al. (2017) found a strong relation between spiral-arm pitch angle and supermassive black-hole mass in spiral galaxies. Other studies have used pitch-angle distributions to investigate spiral-arm formation and winding mechanisms (Hart et al., 2017; Lingard et al., 2021). These applications are presently limited by the cost and methodological heterogeneity of obtaining measurements over very large samples.

An end-to-end regressor could be used as a first-pass catalogue tool. For example, a large imaging survey could be filtered for clearly spiral, moderately inclined systems; the CNN could then produce rapid approximate pitch-angle estimates; and subsets of scientific interest could subsequently be re-measured with Fourier, arm-tracing, or Bayesian methods. Such a two-tier strategy would exploit the speed of deep learning without sacrificing the interpretability of specialised measurement techniques where precision is essential.

Several limitations should be considered when interpreting the reported errors. First, the external validation sample contains only 65 galaxies. It is valuable because the pitch angles were measured independently in the literature and were not used for network optimisation, but it is too small to characterise performance robustly across the full diversity of spiral galaxies. A larger validation set should be stratified by inclination, arm multiplicity, bar strength, flocculence, signal-to-noise ratio, apparent angular size, and pitch angle itself.

Second, the generator assumes an underlying logarithmic spiral with a single pitch angle. Real arms can change pitch with radius or differ from one another (Savchenko & Reshetnikov, 2013; Davis & Hayes, 2014). The present target should therefore be interpreted as an effective global pitch angle consistent with the literature labels used for validation. Extending the generator to radially varying pitch and asymmetric arm geometries would reduce this model mismatch.

Third, the automatic deprojection uses an ellipse fit and a thin-disc approximation. Strong bars, prominent bulges, tidal distortions, and near-edge-on orientations can bias the recovered axis ratio. A future implementation should either exclude unsuitable systems using explicit quality criteria or incorporate independent photometric estimates of disc geometry. It would also be useful to propagate deprojection uncertainty into the final pitch-angle uncertainty.

Fourth, the current study compares the CNN with literature values but does not run 2DFFT, SpArcFiRe, or Spirality on exactly the same 65 images under identical preprocessing. Such an object-by-object benchmark would strengthen any claim about practical competitiveness. The present results support feasibility and demonstrate a clear improvement from geometry-aware preprocessing, but they do not establish superiority over specialised pitch-angle algorithms.

Finally, the current model returns a point estimate only. The tendency toward mean predictions at the extremes shows why predictive uncertainty is essential. Deep ensembles, Monte Carlo dropout, or probabilistic regression could provide uncertainty intervals and enable automated quality control. Calibration should then be evaluated against repeated human or algorithmic measurements rather than against a single reference value.

5 Conclusions

We developed a synthetic-to-real deep-learning pipeline for automated regression of spiral-arm pitch angle from galaxy images. The method combines a controllable logarithmic-spiral generator with exact labels, two-pass rendering under variable inclination, automatic disc deprojection, spatial-frequency enhancement, and two-stage transfer learning with MobileNetV2. On 2,200 synthetic test images, the full pipeline achieved MAE 1.01° and RMSE 1.34° . On 65 real SDSS galaxies with independent literature pitch-angle measurements, the full preprocessing pipeline reduced MAE from 7.07° to 5.83° and RMSE from 10.42° to 7.61° .

These results demonstrate that explicit geometric and image-processing priors can improve the transfer of a CNN trained predominantly on synthetic data. The remaining gap between synthetic and real performance, together with compression of predictions toward the centre of the pitch-angle range, shows that the method is not yet a precision replacement for Fourier, arm-tracing, or template-fitting techniques. Its more immediate value is as a scalable first-pass estimator and as a framework for future hybrid training with larger real labelled samples. The most important next steps are external validation on a substantially larger homogeneous catalogue, direct benchmarking against established pitch-angle codes on identical images, uncertainty quantification, and extension from a single global pitch angle to arm-specific or radially varying predictions.

Acknowledgements. The authors acknowledge the Sloan Digital Sky Survey and the Galaxy Zoo project for providing public astronomical imaging and morphological resources that enabled development of the analysis pipeline. The authors also thank the authors of the published pitch-angle catalogue used for external validation.

Funding. This work supported by the Russian Science Foundation (grant no. 23-71-00016, <https://rscf.ru/project/23-71-00016/>). The research is carried out using the equipment of the shared research facilities of HPC computing resources at Lomonosov Moscow State University.

Data availability. The observational images used in this study are based on public Sloan Digital Sky Survey data. Reference pitch-angle values for the external validation set were obtained from the published catalogue of Yu & Ho (2019). The synthetic images were generated by the software developed for this study.

Code availability. The source code used for synthetic galaxy generation, image preprocessing, model training, and inference is publicly available in the GitHub repository: <https://github.com/Pixeldoom/Automated-Spiral-Arm-Pitch-Angle-Estimation-in-Galaxy-Images-Using-Deep-Learning>.

References

- Abazajian, K.N., Adelman-McCarthy, J.K., Agueros, M.A., et al. 2009, ApJS, 182, 543.
- Barchi, P.H., de Carvalho, R.R., Rosa, R.R., et al. 2020, Astronomy and Computing, 30, 100334. DOI 10.1016/j.ascom.2019.100334.

477 Berrier, J.C., Davis, B.L., Kennefick, D., et al. 2013, ApJ, 769, 132. DOI 10.1088/0004-
 478 637X/769/2/132.

479 Ćiprijanović, A., Lewis, A., Pedro, K., et al. 2023, Machine Learning: Science and
 480 Technology, 4, 025013. DOI 10.1088/2632-2153/acca5f.

481 Díaz-García, S., Salo, H., Knapen, J.H., & Herrera-Endoqui, M. 2019, A&A, 631, A94.
 482 DOI 10.1051/0004-6361/201936000.

483 Davis, D.R., & Hayes, W.B. 2014, ApJ, 790, 87. DOI 10.1088/0004-637X/790/2/87.

484 Davis, B.L., Kennefick, D., Kennefick, J., et al. 2015, ApJL, 802, L13. DOI
 485 10.1088/2041-8205/802/1/L13.

486 Davis, B.L., Berrier, J.C., Shields, D.W., et al. 2012, ApJS, 199, 33.

487 Davis, B.L., Graham, A.W., & Seigar, M.S. 2017, MNRAS, 471, 2187.

488 Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. 2009, in Proc. IEEE
 489 Conf. Computer Vision and Pattern Recognition, 248–255.

490 Dieleman, S., Willett, K.W., & Dambre, J. 2015, MNRAS, 450, 1441. DOI
 491 10.1093/mnras/stv632.

492 Dobbs, C., & Baba, J. 2014, PASA, 31, e035. DOI 10.1017/pasa.2014.31.

493 Domínguez Sánchez, H., Huertas-Company, M., Bernardi, M., Tuccillo, D., & Fischer,
 494 J.L. 2018, MNRAS, 476, 3661. DOI 10.1093/mnras/sty338.

495 D’Onghia, E., Vogelsberger, M., & Hernquist, L. 2013, ApJ, 766, 34. DOI
 496 10.1088/0004-637X/766/1/34.

497 Grand, R.J.J., Kawata, D., & Cropper, M. 2013, A&A, 553, A77. DOI 10.1051/0004-
 498 6361/201321308.

499 Hart, R.E., Bamford, S.P., Keel, W.C., et al. 2017, MNRAS, 472, 2263. DOI
 500 10.1093/mnras/stx2137.

501 Hewitt, I.B., & Treuhardt, P. 2020, MNRAS, 493, 3854. DOI 10.1093/mnras/staa354.

502 Honig, Z.N., & Reid, M.J. 2015, ApJ, 800, 53. DOI 10.1088/0004-637X/800/1/53.

503 Hubble, E.P. 1926, ApJ, 64, 321.

504 Huertas-Company, M., Gravet, R., Cabrera-Vives, G., et al. 2015, ApJS, 221, 8. DOI
 505 10.1088/0067-0049/221/1/8.

506 Ivezić, Ž., Kahn, S.M., Tyson, J.A., et al. 2019, *ApJ*, 873, 111. DOI 10.3847/1538-
507 4357/ab042c.

508 Khrapov, S., Khoperskov, A., & Korchagin, V. 2021, *Galaxies*, 9, 29. DOI
509 10.3390/galaxies9020029.

510 Lee, M.W., Soo, J.Y.H., & Sharoni, S.M.H. 2025, *Astronomy and Computing*, 52,
511 100965. DOI 10.1016/j.ascom.2025.100965.

512 Lin, C.C., & Shu, F.H. 1964, *ApJ*, 140, 646.

513 Lingard, T., Masters, K.L., Krawczyk, C., et al. 2021, *MNRAS*, 504, 3364. DOI
514 10.1093/mnras/stab1072.

515 Lintott, C.J., Schawinski, K., Slosar, A., et al. 2008, *MNRAS*, 389, 1179.

516 Masters, K.L., Lintott, C.J., Hart, R.E., et al. 2019, *MNRAS*, 487, 1808. DOI
517 10.1093/mnras/stz1153.

518 Nair, P.B., & Abraham, R.G. 2010, *ApJS*, 186, 427. DOI 10.1088/0067-
519 0049/186/2/427.

520 Pan, S.J., & Yang, Q. 2010, *IEEE Transactions on Knowledge and Data Engineering*,
521 22, 1345. DOI 10.1109/TKDE.2009.191.

522 Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. 2018, in *Proc.*
523 *IEEE/CVF Conf. Computer Vision and Pattern Recognition*, 4510–4520.

524 Savchenko, S.S., & Reshetnikov, V.P. 2013, *MNRAS*, 436, 1074. DOI
525 10.1093/mnras/stt1627.

526 Seigar, M.S., & James, P.A. 1998, *MNRAS*, 299, 685.

527 Seigar, M.S., Bullock, J.S., Barth, A.J., & Ho, L.C. 2006, *ApJ*, 645, 1012. DOI
528 10.1086/504463.

529 Seigar, M.S., Kennefick, D., Kennefick, J., & Lacy, C.H.S. 2008, *ApJL*, 678, L93. DOI
530 10.1086/588727.

531 Shields, D., Boe, B., Pfountz, C., et al. 2022, *Galaxies*, 10, 100. DOI 10.3390/galax-
532 ies10050100.

533 Shorten, C., & Khoshgoftaar, T.M. 2019, *Journal of Big Data*, 6, 60. DOI
534 10.1186/s40537-019-0197-0.

535 Smith, M.J., & Geach, J.E. 2019, *MNRAS*, 490, 4985. DOI 10.1093/mnras/stz2886.

- 536 Walmsley, M., Lintott, C., Geron, T., et al. 2022, MNRAS, 509, 3966. DOI
537 10.1093/mnras/stab2093.
- 538 Willett, K.W., Lintott, C.J., Bamford, S.P., et al. 2013, MNRAS, 435, 2835.
- 539 Willmott, C.J., & Matsuura, K. 2005, Climate Research, 30, 79.
- 540 York, D.G., Adelman, J., Anderson, J.E., et al. 2000, AJ, 120, 1579.
- 541 Yu, S.-Y., & Ho, L.C. 2019, ApJ, 871, 194. DOI 10.3847/1538-4357/aaf895.