THE UNIVERSITY *of* EDINBURGH
School of Physics
and Astronomy

Cosmological Parameter Inference Using Deep Learning on Simulated Weak Lensing Maps

MPhys Project Report

Diego M. Fernández

Submitted for the 40pt MPhys Project course PHYS11016
19th May 2022

Abstract

In this research, we produce a novel deep learning usage of the *cosmo*-SLICS suite of weak lensing simulations, and analyse model performance in the prediction of cosmological parameters. We aim to improve predictive accuracy on these noise-free maps by exploring various convolutional neural network architectures and hyperparameters, varying the number of layers, learning rates, and number of inputted cosmological parameters. This report concludes by recommending our preferred architecture for future studies, which yielded 98.50% test accuracy on the S_8 parameter. This indicates that deep learning cosmic shear analysis has the potential to outperform the standard two-point weak lensing statistical methods in cosmological parameter constraint.

Supervisors: Professor C. Heymans, Dr. B. Giblin

Acknowledgments

I would like to thank supervisors Dr. Ben Giblin and Prof. Catherine Heymans for their support throughout the project. They provided excellent guidance, answers to technical questions, suggestions for my work, and general help beyond the scope of the project.

Contents

1 Introduction

Cosmological parameter inference – the process of determining specific values that describe the properties of the Universe – has proven an exciting challenge for cosmologists since the late 1990s [1]. Motivated by the confirmation of the accelerating expansion rate of the universe via two separate observational studies of distant supernovae in 1998 by Reiss et al. [2] and Perlmutter et al. [3], the field of cosmology went in a new direction. This discovery proved the existence of the cosmological constant, Λ, which opposes the force of gravity, and is now commonly referred to as 'dark energy' [1]. This lead to the introduction of the Λ-CDM model (also known as the standard cosmological model), relating cold dark matter (CDM) and dark energy. CDM is slow moving and weakly interacts with baryonic ('ordinary') matter via the gravitational force, but it is not currently known to interact via any other forces. The development of this model led to the introduction of six cosmological parameters [1].

The aforementioned 1998 studies calculated coinciding values for cosmological parameters, H_0, the Hubble constant, and Ω_m, the matter density of the universe [1]. The field shifted its focus to constraining the six primary cosmological parameters – a new cosmological research status quo. A variety of complex methods were developed in the following two decades, including parameter constraints via the cosmic microwave background (CMB), improved observational studies of Type Ia supernovae, weak lensing analyses, galaxy clusters, baryon acoustic oscillations (BAO), and combinations of these [4]. Constraints developed by these methods, however, did not always coincide; tensions began to appear in this cosmological model, namely in the constraints of the Hubble parameter, H_0, and the amplitude of matter density fluctuations parameter, S_8, which are further discussed in Section 2.1.2. Cosmologists were left perplexed as to why this was the case. Could the Λ-CDM model be flawed? Could there be something wrong with our understanding of gravity? Could Einstein's general relativity (GR) be imperfect? Could this Universe be a statistical anomaly? Or could we be misinterpreting the early and/or late Universe? All these questions pushed for tighter constraints on the uncertainties' of each of the studies, which surprisingly only exacerbated the tensions [5].

In this project, we investigate weak lensing and take on our own novel approach at constraining cosmological parameters and gaining a better understanding of gravity. Weak lensing refers to the observed elliptical distortion of light emitted from galaxies due to massive structures in our line of sight on a scale of approximately 1%, which is the lower limit of what is referred to as strong lensing [1], as depicted in Figure 1. This method is the conventional method of constraining the S_8 and Ω_m (and sometimes other) cosmological parameters, and is done via statistical analysis of measured galaxy correlations as depicted by cosmological maps from telescope surveys and suites of simulations. The background section of this report details how weak lensing analysis is performed and describes the simulations used to produce our dataset.

The novel aspect of this project is the use of deep learning on the 2019 suite of simulations, *cosmo*-SLICS [7]. Deep learning is a subset of artificial intelligence that aims to emulate the

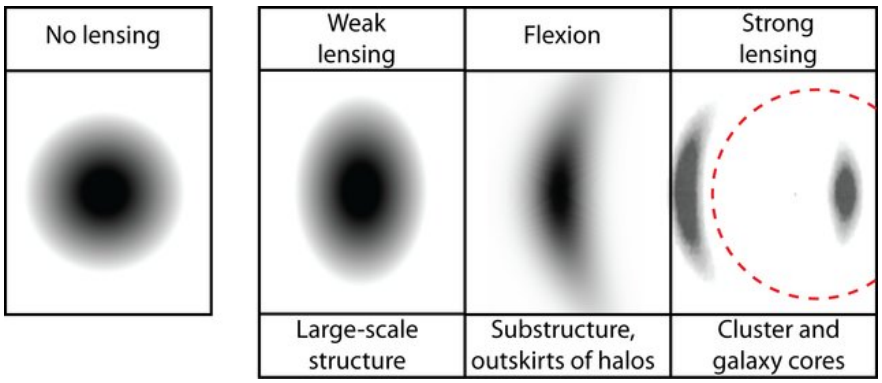


Figure 1. Depiction of weak lensing vs. flexion vs. strong lensing on a perfectly circular image representing a galaxy [6]. Weak lensing distorts the ellipticity of the galaxy, whereas strong lensing can create arcs and/or multiple images of the same galaxy.

brain [8]. It is capable of taking massive amounts of data, on the scale of 10^{10} data points [9], and learning from it to produce predictions on a new dataset. For example, a deep learning model could be given thousands of pictures of cats and dogs with labels for each image. It would then learn the features of cats and dogs via the use of statistics, linear algebra, and other mathematics, and could predict the labels of a new set of images of cats and dogs with a high degree of accuracy [9]. Deep learning is becoming a popular analytical tool and its future uses in cosmology could prove field-changing.

Traditional weak lensing statistical analysis has shown limitations in its ability to constrain cosmological parameters. These limitations arise from the non-Gaussian distribution of matter in the Universe [7]; our statistical methods struggle with such distributions. Research by Fluri et al. 2022 [10] and Fluri et al. 2019 [11], however, has shown that deep learning can be used effectively in the constraint of cosmological parameters using telescope surveys (the KiDS-450 and KiDS-1000 weak lensing maps, respectively). Thus, our aim is to expand upon this deep learning methodology via application on a new simulated dataset. We focus our work on improving the constraints of four cosmological parameters Ω_m, S_8, h, and w_0, defined as the matter density of the universe, the amplitude of matter density fluctuations, the dimensionless reduced Hubble parameter, and the dark energy equation of state, respectively. We do this by, for the first time, applying deep learning to the weak lensing maps produced by the suite of simulations, *cosmo*-SLICS. The deep learning models learn from the majority of the data and then predict on an smaller, unseen subset of the data. This project yields excellent results, with one model predicting the S_8 parameter with 98.50% accuracy, and depicts the potential of deep learning in cosmic shear analysis, whilst suggesting extensions of the project for future work.

2 Background

In this section we discuss the background of the cosmology relevant to the project. In addition, we describe the dataset used and explain the type of deep learning analysis that

was performed.

2.1 Cosmology: wCDM, Cosmological Parameters, and Tensions

In this study, we investigate a set of four cosmological parameters: S_8, Ω_m, h and w_0 (which were defined in the introduction). Here S_8 is dependent on parameters σ_8 and Ω_m by $S_8 \equiv \sigma_8\sqrt{\frac{\Omega_m}{0.3}}$, where σ_8 is the amplitude of the power spectrum [1]. Specifically, we focus on S_8 and Ω_m, which measure the distribution of matter in the Universe, and the density of matter in the Universe, respectively. These two are clearly related [12]. A low value of S_8 suggests a smooth and constant distribution of matter, and a low value of Ω_m suggests there is a relatively low matter density in the Universe [13]. Lensing struggles to differentiate between the two, so studying them together in our deep learning analysis will help depict their relationship. In addition, we discuss the aforementioned tensions in cosmological parameter measurements, namely the S_8 and H_0 tensions.

2.1.1 The wCDM Model

The selected dataset, detailed further in Section 2.2.3, is based in the wCDM cosmological model, which is different from the Λ-CDM model via permitted variations of the w parameter [14]. The w parameter, as referenced in the introduction, is the dark energy equation of state and relates dark energy's pressure, p, to its density ρ in a linear fashion by

$$w = \frac{p}{\rho}. \tag{1}$$

Here, w_0 is treated as a dimensionless parameter. This simplified equation of state essentially shows that when dark energy expands, pressure increases [15]. In the Λ-CDM model, the value of w is fixed to -1, whereas it may vary in the wCDM model and thus affects dark energy and the expansion of the Universe [15].

2.1.2 Tensions in the Standard Cosmological Model

There are two primary disagreements in the standard cosmological model: the Hubble (H_0) tension and the S_8 tension (also referred to as the σ_8 problem) [1]. These arise from discrepancies between different independent observational studies of the Universe and are interconnected. Only a marginal overlap between possible values for these cosmological parameters exists between the independent studies, and this gap shrinks further as accuracy only continues to improve. This suggests the need for a new understanding of the standard cosmological model [1].

2.1.2.1 The Hubble Tension The Hubble tension arises from conflicts in H_0 estimations from Cosmic Distance Ladder (CDL) and Cosmic Microwave Background (CMB) observations.

The first of these measures H_0 in the late Universe – within the last 56 million years – and consists of using extrasolar and extragalactic standard candles [16]. These standard candles are defined as observed sources of light in the Universe with known luminosities, namely Cepheid variable stars (bright pulsating stars), tip of the red giant branch stars (TRGBs), and Type Ia Supernovae [17]. The CDL is made up of steps from nearby standard candles to faraway standard candles, with each standard candle's measurement depending on the prior one. Cepheids and TRBGs are both used as standard candles that separately step up to Type Ia Supernovae [17] [18]. The Cepheid method was developed by Reiss et al. in 1998 [2], whereas the more recent TRGB method was developed by Freedman et al. in 2019 [17]. The most recent estimated values of H_0 from these scientists are $H_0 = 73.04 \pm 1$ km s^{-1} Mpc^{-1} [18] and $H_0 = 69.8 \pm 1.3$ km s^{-1} Mpc^{-1} [17], respectively.

These measurements do not agree with each other, nor with CMB predictions. CMB observations come from the ESA's Planck mission [19], which sought to map the early universe – at 380,000 years old – via high resolution imaging of temperature variations across the early Universe. The first CMB calculation of H_0 assumed the Λ-CDM model and yielded a value of $H_0 = 67.4 \pm 0.5$ km s^{-1} Mpc^{-1} [17]. Bennett et al. 2013 suggests there is a 1% concordance between the CMB and Cepheid measurements, which roughly coincides with the TRGB measurement of $H_0 = 69.6 \pm 0.7$ km s^{-1} Mpc^{-1}, also noting that Baryon Acoustic Oscillation H_0 estimates are in range at $H_0 = 68.76 \pm 0.84$ km s^{-1} Mpc^{-1} [20] [21]. However, as more studies are published, the precision of each measurement only seems to improve, further exacerbating the tension, as can be seen in Figure 2.

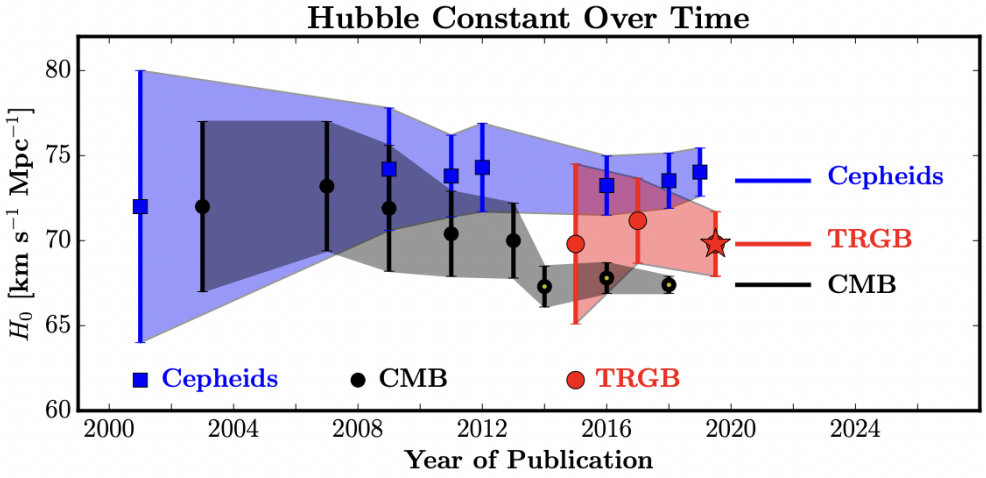


Figure 2. A graphical depiction of the divergence of the Cepheid, TRGB, and CMB measurements of H_0 [17].

2.1.2.2 The S_8 Tension The S_8 Tension emerges in discrepancies of σ_8 and Ω_m from weak lensing and CMB measurements. The values and standard deviations from different

studies are depicted in Figure 3 in which σ_8 is plotted against Ω_m to produce contours of possible values for $S_8 \equiv \sigma_8\sqrt{\Omega_m/0.3}$. As is clear in the figure, the Planck CMB calculation only marginally coincides with the KiDS-1000 gravitational weak lensing analysis (which maps up to 10 billion years in the past), leading to a mere 1% concordance [13]. This suggests that there may be something missing or incorrect in the Λ-CDM model and general relativity, or that unidentified mistakes were made in these studies. An alternative possibility is that the early Universe and late Universe evolved in unexpected ways.

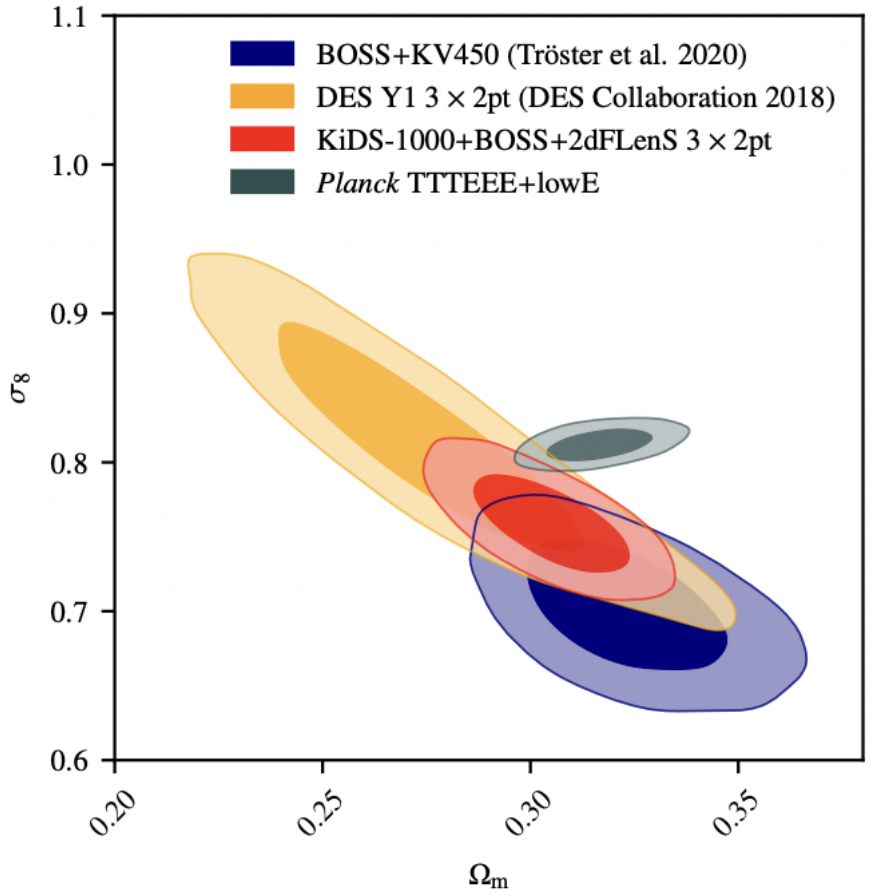


Figure 3. A contour plot of σ_8 vs. Ω_m depicting the tension between 4 different estimates of S_8 [13].

The other two analyses in Figure 3, BOSS+KV450 (Tröster et al. 2020 [22]), and DES Y1 3x 2pt (DES Collaboration 2018) [23], use baryon acoustic oscillations and galaxy clustering, and galaxy-galaxy lensing and weak lensing, respectively, in their constraints of the cosmological parameters. The contour plot only shows two cosmological parameters on the axes, but in reality is a multidimensional parameter space [13]. It is interesting to note that weak lensing analysis cannot differentiate well between high σ_8 and low Ω_m and vice versa, yielding the red contour in the figure [4].

2.1.2.3 Relating the Tensions It is important to note that for both the H_0 and S_8 tensions, there is only a marginal overlap between the conflicting studies' predictions. The uncertainties on the independent studies for each tension decrease every year and reduce their overlap with each other, suggesting a need for an updated cosmological model [17]. The CMB calculations are dependent the Λ-CDM model – an assumption that could lead to potential errors – and focus on the early Universe [1]. On the contrary, CDL and weak lensing measurements are more independent of the model, focusing on observational cosmology, and are based in the more recent Universe [5]. In the CDL calculations, there are limited numbers of standard candle data points on the final step of the ladder – a lack of data which in turn could lead to errors [17]. In terms of weak lensing errors, there are many systematics at play, namely intrinsic alignment, baryonic feedback, and galaxy redshift distributions (further discussed in Section 2.2.2) [4]. These potential flaws suggest a need for further studies into each set of measurements, which this project aims to advance by analysing the *cosmo*-SLICS dataset with deep learning for the first time.

Assuming the independent studies have been performed correctly, and that there are no significant errors, one possible solution could be that the evolution of the early Universe behaved differently from the evolution of the late Universe; this suggests a misunderstanding of general relativity, which is one way of interpreting the tension. Many new cosmological models have been developed in an attempt to reconcile the tensions [1]. Models that address the tensions individually tend to strain the other tension further [5], suggesting that the tensions are intertwined. It is interesting to note that the value of S_8 is lower for weak lensing measurements than for CMB measurements, whereas a higher value of S_8 means a 'clumpier' Universe [24]. For emphasis, we quote *The Dark Universe* by Professor Catherine Heymans, in which she comments that "gravitational lensing has been heralded as the most powerful technique for studying the dark Universe... [but] the issue is that out of all of the four probes in our toolkit, weak lensing is by far the most technologically challenging" [24]. This is a technological challenge we intend to confront in our research.

One solution to the H_0 tension uses the wCDM model, in which w_0 is allowed to vary, as mentioned in Section 2.1.1; however, this only worsens the S_8 tension [25]. Larger and higher resolution future surveys, such as Euclid and SKA, may prove critical in determining whether there is a fault in the studies, or whether something is missing in the standard cosmological model [26]. In this project we expand on this information with our dataset; the simulations are based in the wCDM model [7], and our primary metric of success is the proper calculation of S_8.

2.2 Weak Lensing

Weak lensing is defined by the elliptical distortion of galaxies due to massive structures in our line of sight on a scale of 1%, which is the lower limit of general (or strong) gravitational lensing [1]. The dataset studied in the project consists of simulations of weak lensing maps – images of the sky with these distorted galaxies. We first explore the traditional weak lensing constraints on cosmological parameters using two-point weak lensing functions, then further

discuss the maps and our dataset of choice, *cosmo*-SLICS. Finally, we comment on sources of uncertainty in the understanding of weak lensing maps known as weak lensing systematics.

2.2.1 Two-Point Weak Lensing Functions

We define our baseline weak lensing model using two-point weak lensing functions sensitive to cosmological parameters S_8, σ_8, H_0, and Ω_m (defined in Section 2.1), where $S_8 \equiv \sigma_8\sqrt{\Omega_m/0.3}$ [1]. The functions are particularly sensitive to σ_8 and Ω_m, whose contributions can be cumulatively considered via the S_8 parameter, a topic of importance later on in the development of the project's deep learning models. The aim of these functions is to describe how correlated galaxy shapes are as a function of angular distances, which in practice are typically measured over angular scales of roughly 0.01 to 5 degrees [4]. We consider the depth of space in these functions via the combination of data at different distances, or redshifts, using what are known as tomographic redshift bins [1].

The first in the set of equations is the weak lensing convergence power spectrum, $P_\kappa^{ij}(\ell)$, which includes the matter power spectrum, P_δ, that describes the clustering of matter at different scales. Here, in Equation 2, i and j are the tomographic redshift bins, κ is the convergence, and ℓ (measured in degrees^{-1}) is the angular wave number – the Fourier conjugate of the angular distance (measured in degrees) [27]. This equation is defined by

$$P_\kappa^{ij}(\ell) = \int_0^{\chi_H} d\chi \frac{q_i(\chi)q_j(\chi)}{[f_K(\chi)]^2} P_\delta(\frac{\ell}{f_K(\chi)}, \chi). \tag{2}$$

The integral sums up the total matter over a physical 3D space with distance considered in terms of comoving radial distance, χ, in the limits of lower bound zero and upper bound χ_H, where χ_H is the comoving radial distance from the horizon. $f_K(\chi)$ is the comoving angular diameter distance at χ and $q_i(\chi)$ and $q_j(\chi)$ are the lensing efficiency functions at respective redshift bins i and j [27].

The lensing efficiency function, $q_i(\chi)$, depends on the number of galaxies at a given radial distance within the redshift bin and on the given cosmological values of Ω_m and σ_8 given by

$$q_i(\chi) = \frac{3H_0^2\Omega_m}{2c^2}\frac{f_K(\chi)}{a(\chi)}\int_\chi^{\chi_H} d\chi' n_i(\chi')\frac{f_K(\chi'-\chi)}{f_K(\chi')}. \tag{3}$$

As seen in Equation 3, the lensing efficiency function sums up all the galaxies within the aforementioned bounds of the universe, with $n_i(\chi') = n(\chi)d\chi$ representing the number of galaxies there are in the given redshift bin, normalised such that $\int_0^{\chi_H} n(\chi)d\chi = 1$. The equation's dimensions are then scaled by scale factor $a(\chi)$ for a given value of χ. As usual, H_0 is the Hubble constant (the current expansion rate of the Universe), Ω_m is the matter density of the universe, and c is the speed of light in a vacuum [27]. The lensing convergence power spectrum is determined from the lensing efficiency function [27] using the Limber

approximation, which relates the appearance of galaxy spacing in their projections to their spatial separations (the curious reader should see [28] for further details).

Given these equations, we can now define the theoretical and observed tomographic angular two-point shear correlation functions $\xi_{\pm}^{ij}(\vartheta)$ and $\hat{\xi}_{\pm}^{ij}(\vartheta)$, respectively, in terms of angular separation ϑ [27]. Provided a Gaussian distribution of matter density in the Universe, $\xi_{\pm}^{ij}(\vartheta)$ would capture all the cosmological information necessary to minimise constraints. The Universe does not, however, have a perfectly Gaussian matter distribution, and thus the constraints are not optimal [1].

Using the lensing convergence power spectrum, $P_{\kappa}^{ij}(\ell)$, we define the theoretical tomographic angular two-point shear correlation function, $\xi_{\pm}^{ij}(\vartheta)$, in terms of the angular separation ϑ [27] and the first and fourth order Bessel functions $J_{0,4}(\ell\vartheta)$, which are beyond the scope of this research, but are relevant to their respective shear correlation functions [29] which are defined as

$$\xi_{\pm}^{ij}(\vartheta) = \frac{1}{2\pi}\int d\ell\ell P_{\kappa}^{ij}(\ell) J_{0,4}(\ell\vartheta). \tag{4}$$

We thus look at the observed tomographic angular shear two-point correlation function, $\hat{\xi}_{\pm}^{ij}(\vartheta)$, associating two galaxies at angular positions a and b, with separation $|\vec{x_a} - \vec{x_b}|$ in interval $\Delta\theta$ to θ. This can be estimated by Equation 5 where w_a and w_b are the weights defining how well galaxy ellipticities are measured, and is then compared to the theoretical model in Equation 4 [27]. The weights give increased or decreased value to their respective galaxies to account for the quality of their measurement. Here, ϵ_t and ϵ_x are the tangential and cross components of the ellipticity of the galaxies in tomographic redshift bins i and j, which allow for the comparison of galaxies closer to and further from Earth. This equation multiplies and adds (or subtracts) the ellipticities of galaxies, giving two different shear correlation functions ξ_{+}^{ij} and ξ_{-}^{ij} by

$$\hat{\xi}_{\pm}^{ij}(\vartheta) = \frac{\Sigma_{ab} w_a w_b [\epsilon_t^i(\vec{x}_a)\epsilon_t^j(\vec{x}_b) \pm \epsilon_{\times}^i(\vec{x}_a)\epsilon_{\times}^j(\vec{x}_b)]}{\Sigma_{ab} w_a w_b}. \tag{5}$$

These equations help constrain the relevant cosmological parameters via the shear correlation function covariance matrix. This covariance matrix contains the errors in measurement, (the shear correlation function) and information on how correlated the measurement is on different angular scales (see [7] for further information). The paper notes that analysis has been done to detail the aforementioned covariance, which is beneficial to use in the two-point correlation functions [7].

2.2.2 Weak Lensing Systematics

Weak lensing systematics are significant causes of uncertainty when constraining cosmological parameters. For completeness, we note and define three of the main systematics: intrinsic

alignments, baryonic feedback and galaxy redshift distribution [4]. Though they are not considered in the analysis of our dataset, these must be examined if our methods are applied to non-simulated data. Systematics are present in real data, though most can be dealt with by data manipulation and calibration [7].

2.2.2.1 Intrinsic Alignments Intrinsic alignments manifest in the form of overlapping galaxies that are not the result of gravitational lensing. One manifestation of this systematic lies in galaxies that are too close to each other in terms of distance from Earth. This could lead to incorrect estimates (typically over-estimates) and biased results of galaxy correlations, the equations of which have been discussed in Section 2.2.1 [4].

2.2.2.2 Baryonic Feedback The interactions of baryonic ('ordinary') matter on large scales (e.g. from Earth to faraway galaxies) yield challenges in weak lensing analysis. Luminous baryonic matter can lead to increased magnification and maps with multiple images of the same galaxy – a challenge that can be difficult to mitigate [30]. Electromagnetic interactions between baryonic matter and the light from distant galaxies cause this problem, and can impact the shear correlation function at small angular separations, thus affecting calculations of the correlation of galaxies [4].

2.2.2.3 Galaxy Redshift Distributions It is imperative in weak lensing analysis that images be distributed in redshift bins. Bin width and mean redshift have the most influence on the correlation calculations, especially impacting the shear correlation function, due to its dependence on the lensing efficiency function, seen in Equation 3. The challenge lies in determining the redshifts of galaxies; if they are placed in incorrect bins, the correlations can be over- or under-estimated. This systematic is not impacted by the effect of gravitational lensing and continues to pose a challenge to cosmologists [4].

2.2.3 The *cosmo*-SLICS Maps

The suite of simulations used in this study, *cosmo*-SLICS, consists of N-body simulations of 25 wCDM cosmological models, varying Ω_m, S_8, h, and w_0 within reasonable range of the Λ-CDM expected cosmology, alongside a fiducial cosmology – the cosmology representing the standard Λ-CDM model, with a fixed value of $w_0 = -1$. The dataset is made up of 5 tomographic redshift bins: z = 0.1-0.3, 0.3-0.5, 0.5-0.7, 0.7-0.9, and 0.9-1.1 [7]. This redshift bin distribution matches the distribution found in the KiDS-1000 data release, which allows for broader future use of the developed deep learning models [31].

The simulations produced cosmological maps for the 26 unique cosmologies in the form of 128x128 pixel arrays distributed across a 5 degree by 5 degree space for 5 redshift bins, with unit-less values roughly between -1 and +1 at each point [7]. These simulations produce Kappa maps – maps that use the shear signal to infer projected matter density onto a 2D plane along the line of sight. Kappa maps differ from shear (or ellipticity) maps in that

they cannot be directly observed, but instead must be calculated from shear maps via a set of Fourier transformations [32]. They depict under-densities as negative values and over-densities as positive values; in Figure 4, the bright spots represent the over-densities and the dark spaces represent cosmic voids [7].

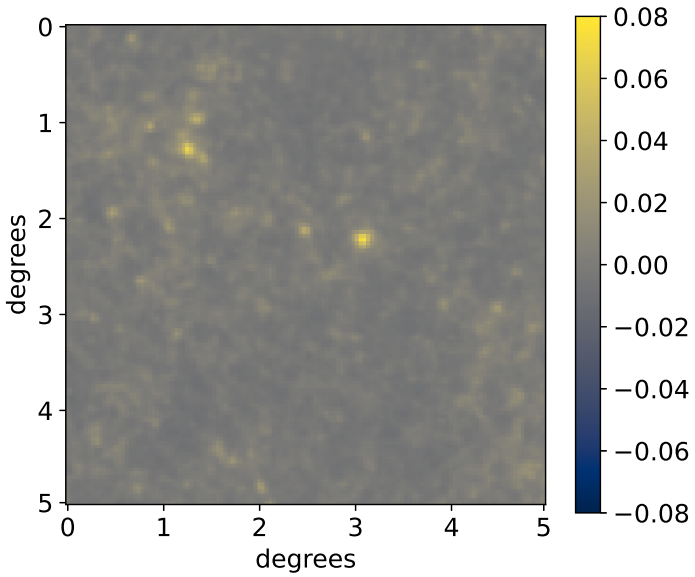
Upon two-point weak lensing analysis (described in Section 2.2.1), and the use of shear correlation function covariance matrices, the simulations met the Harnois-Déraps et al. 2019 [7] criteria of being within 1-2% of expected accuracy in much of the parameter range, citing diminishing returns in the further use of computing resources for reduction of accuracy. The resulting dataset, designed for use in weak lensing analysis, was thus determined effective in performing cosmological parameter inference on the manipulated cosmological parameters as well as the additional σ_8, Ω_c, and Ω_Λ, where Ω_c is the matter density of relativistic particles [7].

The dataset has been used extensively in weak lensing analyses, particularly with machine learning. Some notable works include Harnois-Déraps et al. 2021 [33], Heydenreich et al. 2021 [34], Davies et al. 2021 [35], and Davies et al. 2022 [36]. Despite their use in machine learning and more traditional weak lensing analyses, the maps have never been analysed with deep learning – a novel challenge tackled in this investigation.

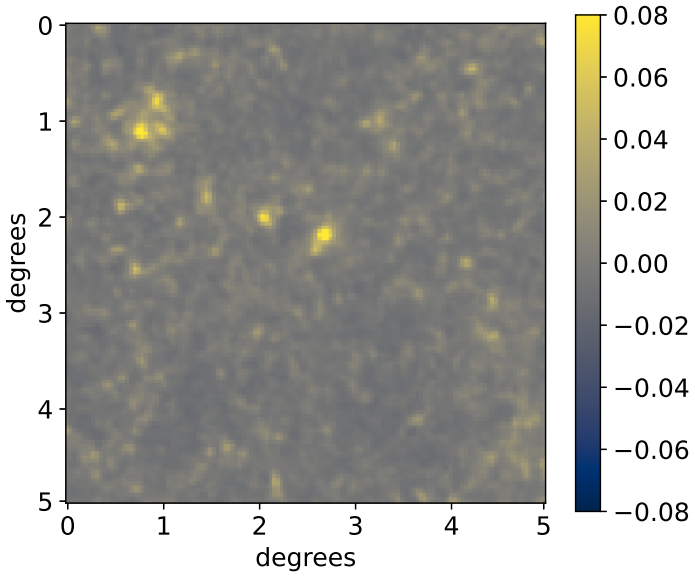
Other studies have attempted deep learning analysis on survey datasets, namely the Kilo-Degree Survey (KiDS), a weak lensing shear tomography survey, used in Fluri et al. 2022 [10] and Fluri et al. 2019 [11] with the KiDS-1000 and KiDS-450 releases, respectively. These datasets contrast with our selection of simulated maps via their inclusion of noise, primarily in the form of varying galaxy ellipticity. The *cosmo*-SLICS maps were developed with no ellipticity; a consistent, intrinsic circular galaxy shape allows for a quicker and more accurate analysis but does not perfectly represent the exact conditions that would be expected in analysis of true survey maps. The focus of this project, however, was to find an optimal convolutional neural network architecture under the restriction of limited computational resources, thus suiting the dataset well. In short, this project takes on the novel task of training deep learning models on *cosmo*-SLICS maps – a dataset never before used in deep learning analysis.

2.3 Deep Learning: Convolutional Neural Networks

Deep learning is a subset of machine learning (ML) – in itself a subset of artificial intelligence (AI) – that aims to emulate the functionality of the brain [9]. With a combination of linear algebra, statistics, and other mathematics, deep learning models can learn from large datasets in a manner more efficient, and often more effective, than traditional hands-on data analytical methods [8]. Deep learning is capable of handling unstructured data (i.e. image pixel arrays, blocks of text, etc.) with the objective of using one dataset to predict information from another dataset. Said dataset can be learned from, whether it uses target labels or no labels at all, referred to as supervised and unsupervised learning, respectively [9]. This project aims to develop a highly accurate predictive model for cosmological parameter constraints on the *cosmo*-SLICS dataset, using a subset of deep learning known as the convolutional neural network (CNN) [9].



(a) Minimum S_8, where $S_8 = 0.6101$.



(b) Maximum S_8, where $S_8 = 0.8947$.

Figure 4. The figures depict weak lensing convergence maps from the *cosmo*-SLICS dataset in redshift bin $0.5 < z < 0.7$. This allows for comparison of low and high S_8 valued maps. The bright, yellow spots represent over-densities and the darker spots represent cosmic voids. Figure 4a clearly has fewer bright spots, which coincides with the nature of a lower S_8 value – less clustering. Figure 4b, on the other hand has more clustering as depicted by the large bright spots.

2.3.1 CNN Architecture

Convolutional neural networks take a dataset (in our case a collection of 2D images with a set of output values as labels), and pass it through a series of layers that perform convolutions – specialised linear operations – on the values in the pixel array [9]. In between convolutions or series of convolutions, the typical CNN will pass the data to a pooling layer, where the data is 'pooled' together to reduce the size of the array. This can be done in many ways, but often consists of taking the mean, maximum, or minimum value of each set of square arrays of a certain size [9]. In the case of the CNN developed in this project, our pooling layer gets the maximum value of each 2x2 square in the 128x128 and 64x64 pixel arrays after the 3rd (Conv-3) and 6th (Conv-7) convolutions, respectively, and shrinks them down to 64x64 and 32x32 grids, respectively. This can be noted in the architecture of our simplest CNN, seen here in Figure 5. At the end of the CNN lies the fully connected layer, which flattens the values down to a 1D array in a linear fashion [9]. After a selected number of epochs – the number of times all the training data passes through the training process – the parameters have been trained and then can be used to predict on an unseen dataset that determines the accuracy of the CNN [9].

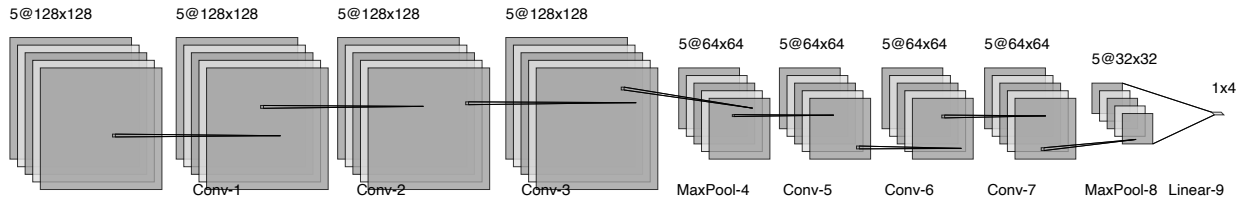


Figure 5. The architecture of the CNN with six convolutional layers and two pooling layers with a ReLU activation function for each. The final layer consists of a fully connected linear regression layer. This architecture corresponds to the simplest architecture used in our model analysis.
This visualisation was constructed using `http://alexlenail.me/NN-SVG/LeNet.html`.

2.3.2 CNN Hyperparameters

Many parameters and hyperparameters of a CNN can be fine-tuned in order to improve its accuracy and loss. These include, but are not limited to: the number of layers (pooling, convolutional, and fully connected), the size of the dataset, the size of the convolutional filter, the size of the pooling square, the batch size – the amount of images that are passed to the CNN in the training process at one time – and the learning rate, which determines how quickly the weights and biases are allowed to change. This non-exhaustive list depicts the extent to which many features can be manipulated in the optimisation of CNN performance. In the case of this project, we focus on three hyperparameters: the number of layers, the learning rate, and the amount of cosmological parameters to be learned, detailed further in Section 3.1.

2.3.3 CNN Model Selection

In this report, we base our model selection on the efficient and commonly used 'early stopping' method, which saves the CNN weights right before they are overfit on the training set [8], or when they are most effective in predicting on the test set. A typical classification CNN is overfit when the accuracy on the test set begins to decrease, the test loss curve begins to increase, and/or the loss on the training set goes to zero as the number of epochs increases [8]. Overfitting, and its converse, underfitting, are further discussed in the analysis of the CNNs produced. An ideal CNN is neither overfit nor underfit when the weights are saved; this generally ensures the best possible performance on the test dataset.

There are many additional scores and metrics that can be used to evaluate the performance of a CNN, as detailed in Alzubaidi et al. 2021 [8], but the majority are challenging to apply to CNNs with regression – CNNs that aim to predict specific values instead of simply classifying the data. The typical classification accuracy calculation is given by

$$\text{Accuracy}_{\text{classification}} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}}. \quad (6)$$

Here, the sum of 'True Positive' and 'True Negative' is the number of correctly classified test samples, and the sum of 'False Positive' and 'False Negative' is the number of incorrectly classified samples. In other words, the accuracy is equal to the amount of correct predictions over the total number of predictions [8].

In our work, we calculate accuracy in two different forms. The first uses a zero-based accuracy metric that includes information about whether the prediction is an over-estimate or an under-estimate, as can be seen in Equation 7. This modified accuracy metric is used for plotting the accuracy of individual cosmologies for each cosmological parameter later on in the methodology. A value of zero represents perfect accuracy, a positive value represents an over-estimate, and a negative value represents and under-estimate, given by

$$\text{Accuracy}_{\text{modified}} = \frac{\text{Predicted Value} - \text{True Value}}{|\text{True Value}|}. \quad (7)$$

The more traditional accuracy metric, however, would be calculated via Equation 8. Here we sum all the cosmologies' modified accuracies at each epoch and divide by six, the number of cosmologies in the test set, to get the mean modified accuracy. We then subtract one by this value to produce the traditional accuracy metric. This accuracy metric is later used to determine the most effective CNN by picking the single maximum value across all the epochs; the maximum value would thus suggest the early stopping epoch. Here, a value of one represents perfect accuracy and a value of zero represents poor accuracy by

$$\text{Accuracy}_{\text{traditional}} = 1 - \frac{\Sigma_{n=1}^{6} |\text{Accuracy(0 based)}|}{6}. \quad (8)$$

Multiplying the traditional accuracy by 100 would yield the accuracy as a percentage, where 100% means a perfect prediction, allowing for simple comparison of model accuracies, as is further discussed in the results section.

3 Methodology

This project focuses on constraining four cosmological parameters (Ω_m, S_8, h, and w_0) via the novel use of deep learning on the *cosmo*-SLICS-developed weak lensing maps. Specifically, the selected models consist of 18 convolutional neural networks (CNNs). Our aim is to compare variations in CNN architecture and hyperparameter selection to see how they affect the network's ability to learn cosmological information from weak lensing simulations. As mentioned previously, the non-Gaussianity of the matter distribution of the Universe suggests that current statistical weak lensing methods are not powerful enough to handle all the challenges in cosmological parameter inference [7], and we aim to use CNNs on weak lensing maps as an alternative solution.

Convolutional Neural Networks (CNNs) have proven effective in image classification tasks. Here we take it one step further and use CNNs with regression to develop a model that can estimate our cosmological parameters of interest with the overall objective of developing the best possible model by tuning a set of parameters and hyperparameters. In particular, we aim to further confirm the relative importance of the parameters S_8 and Ω_m by isolating them and training a CNN with just the two. This idea was studied in Tröster et al. 2021 [12], where they showed that h and w_0 had less impact on weak lensing correlation functions than S_8 and Ω_m. In using just the two parameters, we aim to develop a model with higher accuracy on the test set, despite having less cosmological information. We note in the four parameter CNN training process that none of the parameters have been given any additional weighting, and thus, importantly, the CNN does not have any prior bias in favour of S_8 and Ω_m.

3.1 Model Architecture and Hyperparameter Selection

With the objective of developing the best predicting CNN in mind, we decided to tune three paramaters/hyperparameters. The first of these involves the structure of the model architecture: we wanted to choose between six, eight, and ten convolutional layers. Early tests suggested that fewer layers had diverging accuracy, and that more layers took too long to run (further studies would ideally have the benefit of more computational time). We also chose to vary the learning rates of the models, and included two pooling layers after the third and sixth convolutions; graphic and tabular representations can be seen in Figure 5 and Table 1, respectively. Fluri et al. 2019 used a learning rate of 0.5×10^{-5}, but used a more complex architecture than ours alongside noisier maps, in which the cosmological features are less dominant, which could affect how quickly they would have wanted their weights to change [11]. To compensate for the relative simplicity of our CNN architecture, we selected

smaller learning rates to be trained over a period of 5000 epochs and decided to compare between 1 x 10^{-6}, 0.5 x 10^{-6}, and 1 x 10^{-7}. Additionally, we used their suggested activation function, the rectified linear activation function (ReLU), which outputs non-negative values by selecting the maximum of zero and the inputted parameter [8]. If the inputted parameter, x, is negative, then zero is returned, otherwise, the inputted parameter is returned by

$$f(x)_{ReLU} = max(0, x). \tag{9}$$

Finally, and most importantly, we decided to vary the number of cosmological parameters used in training. S_8 and Ω_m have a more significant impact on weak lensing correlation functions than h and w_0 [12]. Accordingly, we implement a novel analysis via the comparison of training our CNNs on all four of the cosmological parameters and just the two, S_8 and Ω_m.

Table 1. A table depicting the output shape and number of parameters for each layer in the simplest six convolutional layer CNN. This table corresponds to the architecture in Figure 5.

Layer	Output Shape	# of Params (weights)
Conv2d-1	[-1, 5, 128, 128]	230
Conv2d-2	[-1, 5, 128, 128]	230
Conv2d-3	[-1, 5, 128, 128]	230
MaxPool2d-4	[-1, 5, 64, 64]	0
Conv2d-1	[-1, 5, 64, 64]	230
Conv2d-1	[-1, 5, 64, 64]	230
Conv2d-1	[-1, 5, 64, 64]	230
MaxPool2d-4	[-1, 5, 32, 32]	0
Linear-9	[-1, 4]	20,484

The result of using three different layer counts, three learning rates, and two selections of cosmological parameter counts yields a total of 18 CNNs, which can be challenging to compare without a specific metric. Though model selection is not typically based solely on accuracy or loss, due to the complicated nature of comparing 18 models, we select the best model using the traditional accuracy metric in Equation 8. For completeness, we have also produced loss plots that can be used to determine whether the CNN is learning correctly, thus producing a more detailed qualitative analysis of the models.

3.2 Data Management

3.2.1 Train vs. Test Split

The data is split according to proximity of *cosmo*-SLICS cosmologies to the Λ-CDM model. The test data contains half the data from the fiducial cosmology and half the data from the

other five most expected cosmologies in regard to the standard model – a total of 576 maps with five redshifts for four cosmological parameters at six different cosmologies. The rest of the data is put into the training set, with the aim of being representative of a variety of cosmologies within the wCDM model. Our test data acts as our validation set; accuracy and loss are calculated throughout the training process to be plotted afterwards [9]. This allows for fine-tuning of the hyperparameters upon the completion of the training process.

3.2.2 Storing Model Weights and Predictions

Weights and predictions are stored every two epochs (to save memory) in the form of pytorch [37] and csv files, respectively. The weight files can be easily used by running a Python [38] script that predicts on a new (or old) dataset. The csv file can be downloaded into a pandas DataFrame [39] for straightforward data access. The file includes, at every two epochs, values for the predictions, standard deviations, and true cosmologies of each of the four cosmological parameters for each of the six cosmologies, as well as the overall training loss and test loss, and the mean S_8 test accuracy. From this point, the best average accuracy and minimum training loss can be calculated.

3.3 Final Model Selection

We determine the best two and four cosmological parameter CNNs using the accuracy of the S_8 cosmological parameter, as justified in Section 3.1. For this analysis, we have selected the models with the best S_8 accuracy using the traditional accuracy metric demonstrated in Equation 8. Comparison of the S_8 accuracies for the two and four cosmological parameter CNNs can be found in the results section, but we will use the best peak S_8 accuracy models here in our analysis.

3.3.1 Four Cosmological Parameter Analysis

The best S_8 accuracy four cosmological parameter model consists of six layers and a learning rate of $0.5\text{x}10^{-6}$ with a peak S_8 accuracy of 97.07% at the 90th epoch. We begin by producing the plots for loss, accuracy, and predictions/truth values vs. the number of epochs. These can be seen here in Figure 6, where the black lines depict the best accuracies of each cosmological parameter.

We note in this figure that the peak accuracy for S_8 occurs before the test loss curve 'elbow' – the point at which the loss curve begins to flatten rapidly [40]. This indicates that the best S_8 accuracy does not coincide with the model's best overall fit. Instead, it appears that the model has fit S_8 (and Ω_m) prior to w_0 and h, which are likely the sources of higher loss in the loss curve (because the loss curve combines the loss of all the parameters, and not just each individual one). The accuracy values beyond epoch 90 do not get any better, suggesting that the CNN is sacrificing its predictions on S_8 for better predictions on other parameters.

Thus, the weight of S_8 (and, not surprisingly, Ω_m) may be heavier in the training process than the other parameters, making it take more epochs to fit all the parameters more evenly. Similarly, the Ω_m accuracy is also toward the final epochs, depicting the close relationship between the two parameters (we note that S_8 is dependent on Ω_m).

Importantly, we have not given the neural network any sense of which parameters to prioritise, though this could be done in future works. We note, by looking at the true/predicted values plot that some cosmologies' predictions approach their true values as the number of epochs approaches 5000, but not all of them do. For example, the blue and purple lines for cosmologies '00' and '18', respectively, continue in the direction of their true values as the number of epochs increases, but the dark green line – cosmology '07' – moves away from its true value. It is possible that further training (on the scale of tens of thousands of epochs) may improve the model's accuracy, but the limited time of the project did not allow for these tests.

Though not depicted here, nearly half of the four parameter CNNs had their best S_8 accuracies prior to the end of the test loss 'elbow'. This further stresses the importance of S_8 in relation to the other parameters. We continued examining this by looking at the predictions on S_8 produced by corresponding CNNs trained only on the S_8 and Ω_m parameter data in the following section.

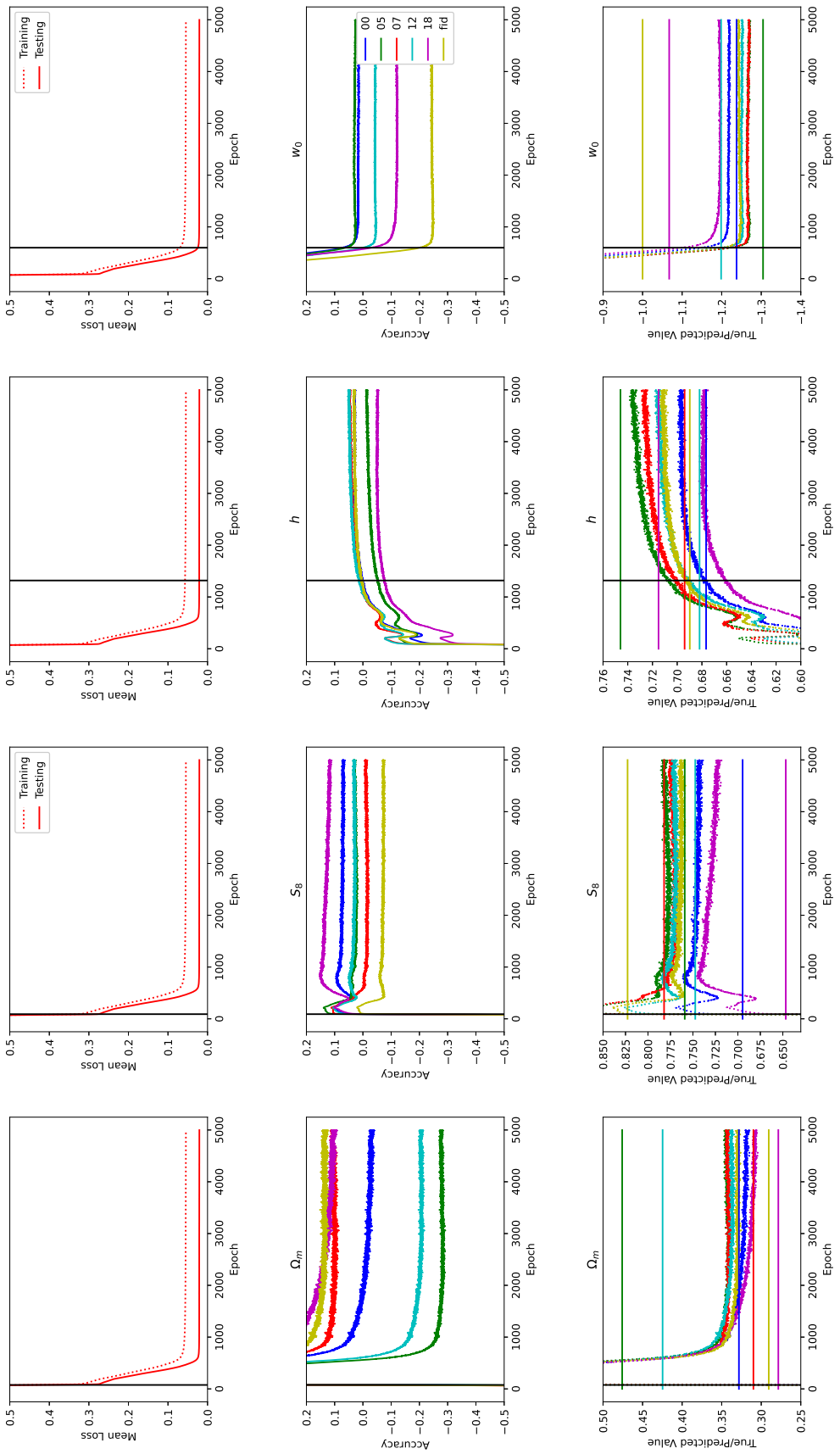


Figure 6. Mean loss, accuracy, and true/predicted value vs. epoch plots for the four cosmological parameter CNN with the best S_8 mean accuracy. The best accuracy of S_8 is seen prior to the elbow of the loss curve, indicating that as the number of epochs goes on, the model sacrifices accuracy on S_8 (and, Ω_m) for improvements in h and w_0.

3.3.2 Two Cosmological Parameter Analysis

The best S_8 accuracy two cosmological parameter model had ten layers and a learning rate of $1\text{x}10^{-6}$ with a peak S_8 accuracy of 98.50% at the 4,956th epoch. Again, we produce the plots for loss, accuracy, and predictions/truth values vs. the number of epochs, seen here in Figure 7, in which the black lines depicts the best accuracy for each parameter.

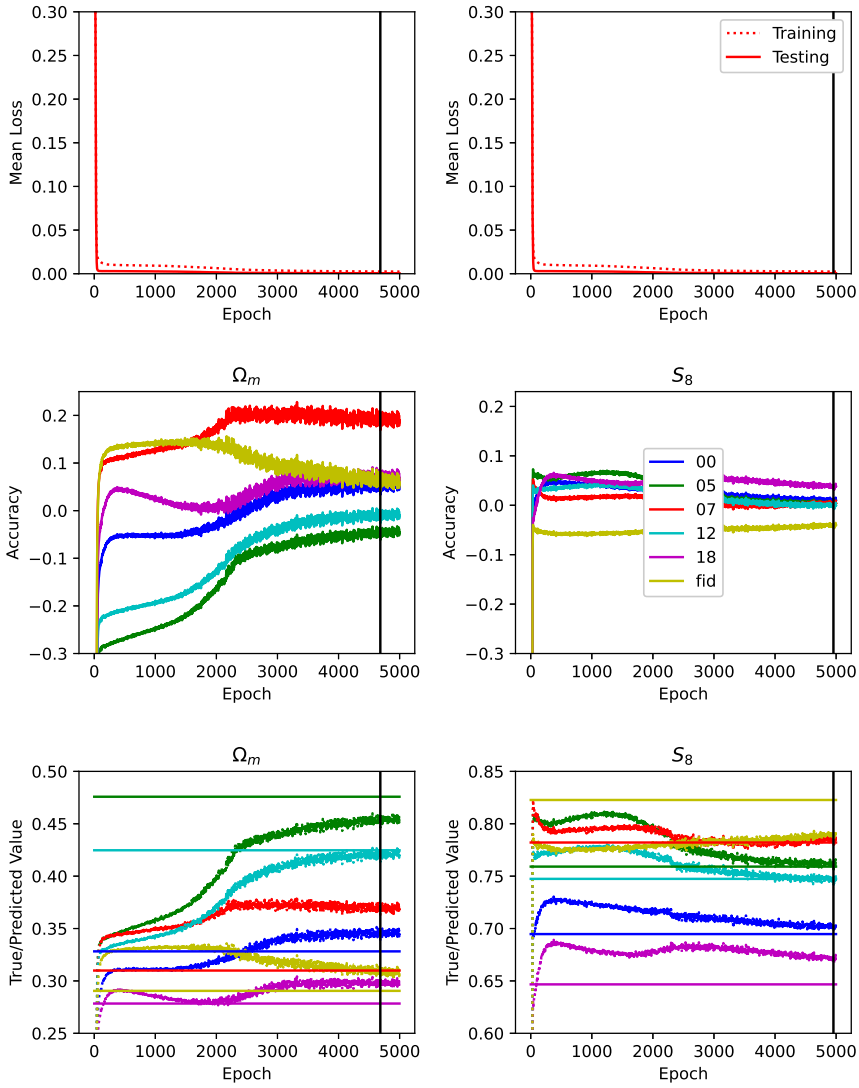


Figure 7. Mean loss, accuracy, and true/predicted value vs. epoch plots for the two cosmological parameter CNN with the best S_8 mean accuracy. The best accuracy of S_8 is seen toward the end of the epochs, suggesting that further training could improve the model's predictive abilities.

The best S_8 accuracy line in these plots differs greatly in position from the four parameter best model; it is essentially at the end of the training period, suggesting that more training

(on the scale of tens of thousands of epochs) might further improve the accuracy. Similarly, the Ω_m accuracy is also toward the final epochs, depicting the close relationship between the two parameters as was also noted in the four parameter analysis.

When looking at the true/predicted value plots, it is interesting to note that the predictions for some of the cosmologies approach their true values, whereas others diverge. This is particularly noticeable on the Ω_m plot, in which the green and cyan cosmologies nearly match the true values toward the final epochs. In similar fashion to the four parameter model, the purple and blue lines tend toward their true values in the S_8 plot, but in this case, the green line does as well. However, for S_8, they generally tend toward the center of the true parameter value space – a trend noted in many of the plots not shown here. This is important to note because it shows that the model has likely learned some cosmologies better than others.

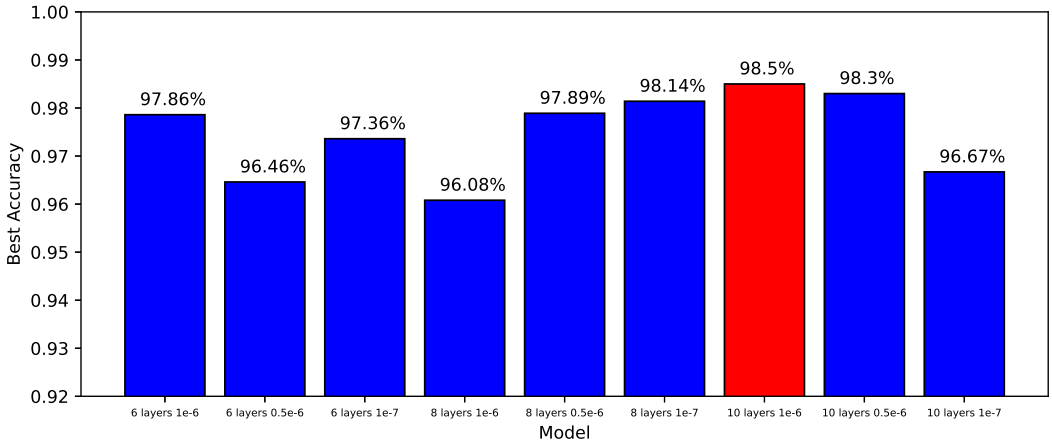
4 Results and Discussion

4.1 Comparing Two and Four Parameter Model Performance

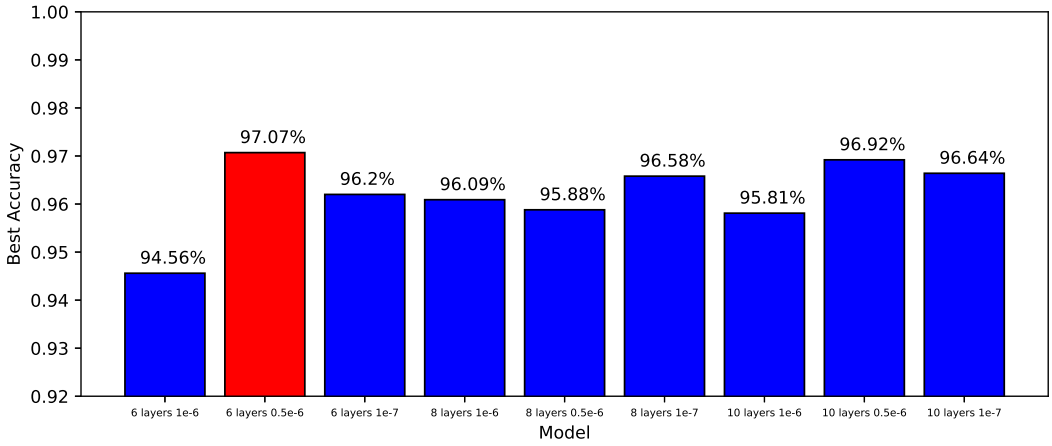
Overall, each model had excellent performance, with a maximum of 98.50% accuracy (on a two parameter model) and a minimum of 94.56% accuracy (on a four parameter model) for the S_8 cosmological parameter. The mean test accuracy of the nine CNNs for each of the two and four parameter models was 97.47% and 96.19% on S_8, respectively. We can see this difference clearly in the bar charts in Figure 8. Though a 1.28% discrepancy may not seem significant, it is worth noting – the two parameter models performed better in all but two of the nine models.

In general, across the plots and all the parameters, as the number of epochs increased in training, the predictions tended to converge near the center of each parameter value space. This explains the consistently high accuracy plots across all the cosmological parameters and their cosmologies. The *cosmo*-SLICS dataset does not have the broadest range of parameter values, so most predictions within the minimum and the maximum of each cosmological parameter value could yield a high accuracy. For example, the minimum training S_8 value is 0.6101 and the maximum is 0.8947. Any accuracy estimate on a prediction between these values would be 68% or better, and thus this is deemed the baseline accuracy of the models, equivalent to a random guess between the values.

Unlike the four parameter CNNs, which had three models with the best S_8 accuracy prior to the test loss 'elbow', only one of the two parameter CNNs had the best S_8 accuracy prior to the test loss 'elbow'. This further supports the insight that S_8 and Ω_m have a greater influence on weak lensing analysis than h and w_0. We also note that the non-smooth accuracy curves suggest the learning rate is too large; the model is over-shooting the prediction by updating its weights too quickly. As a result, even smaller learning rates should be applied in future testing. Finally, we comment that it is likely the models have learned some cosmologies better than others; this is determined from the predicted/true vs. epoch plots, in which some cosmologies' predictions approach their true values.



(a) Two Parameters



(b) Four Parameters

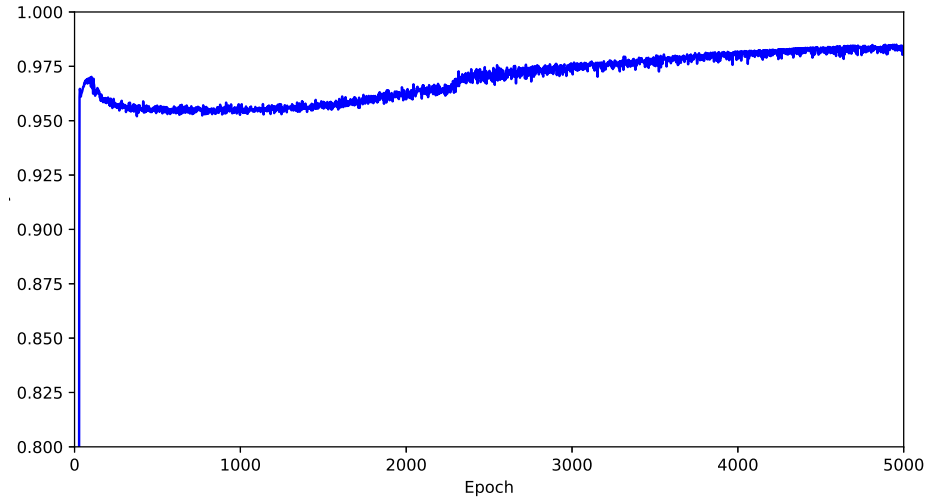
Figure 8. The figures depict bar charts of the best average S_8 accuracies for each of the models. The best performing model is marked by a red bar. Here the best performing model for two parameters is the ten layer one with a learning rate of $1\text{x}10^{-6}$ and a peak S_8 accuracy of 98.50%, whereas the best four parameter model is the six layer one with a learning rate of $0.5\text{x}10^{-6}$ and a peak S_8 accuracy of 97.07%

4.1.1 Mean S_8 Accuracy vs. Epoch of the Best Two and Four Parameter Models

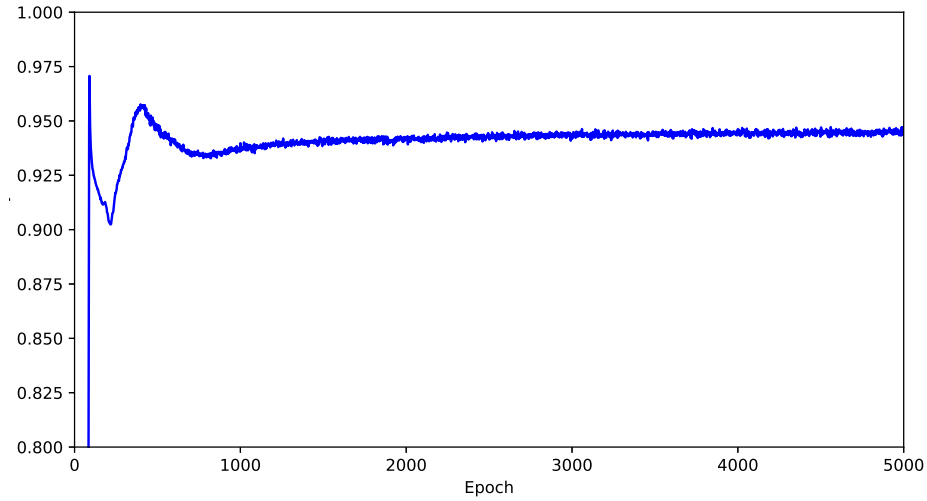
We see in Figure 9a and Figure 9b the mean accuracy vs. epoch plots of the best S_8 predictions for the two and four parameter CNNs, respectively.

Beginning with Figure 9a, we note that despite the initial peak in S_8 accuracy, the accuracy in later epochs actually continues to rise. This plot suggests further training could yield an even higher accuracy.

We note in Figure 9b that the initial spike in S_8 accuracy is the maximum across all the



(a) S_8 Mean Accuracy vs. Epoch for Two Parameters



(b) S_8 Mean Accuracy vs. Epoch for Four Parameters

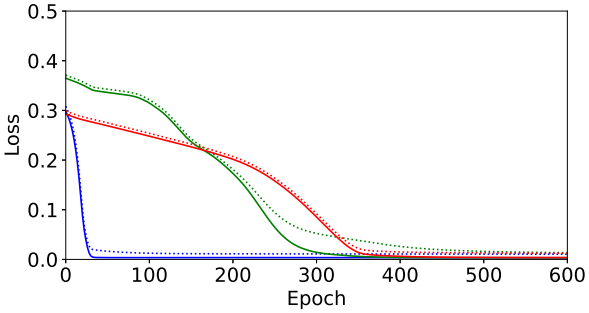
Figure 9. These figures depict the mean S_8 accuracies across the six test cosmologies, plotted against the number of epochs, for the best S_8 accuracy models of the two and four parameter CNNs. As mentioned in section 3.3.2, the peak S_8 accuracy for Figure 9a is clearly toward the final epochs, and seems to be rising further. Figure 9b, depicts a peak S_8 accuracy at epoch 90, as discussed in Section 3.3.1, with all subsequent accuracies below that value; the accuracies appear to plateau.

epochs. After some fluctuations, the curve begins to rise steadily, but appears to plateau well short of the peak S_8 accuracy. This suggests that further training will not improve the model and that it has sacrificed the accuracy of S_8 for other cosmological parameters.

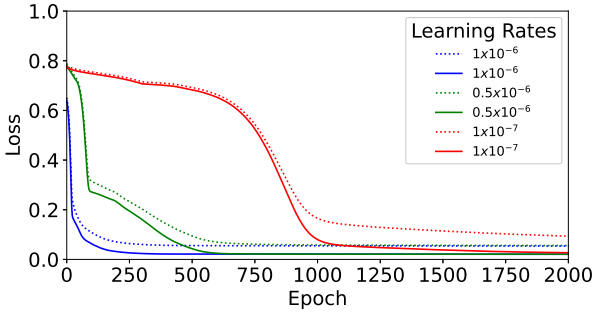
The above comments mark clear differences in the plots, and suggest that more training is better for the two parameter model than for the four parameter model. Despite this, both models have a very high accuracy for S_8 and reach these values quickly, spiking within the

first 100 epochs each.

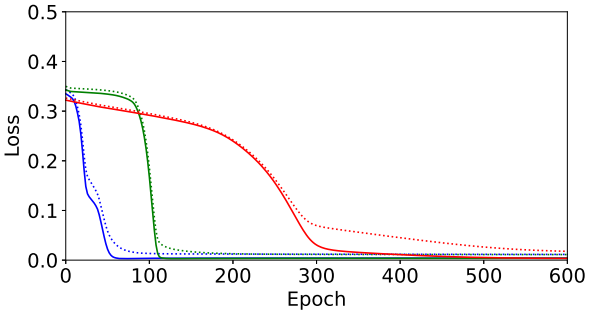
4.2 Learning Rate Analysis: Reviewing the Loss Curves



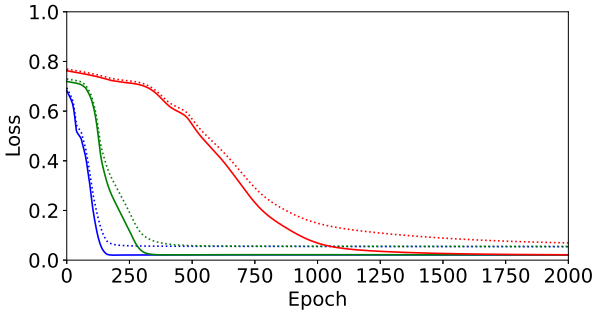
(a) Two Parameters with Six Layers



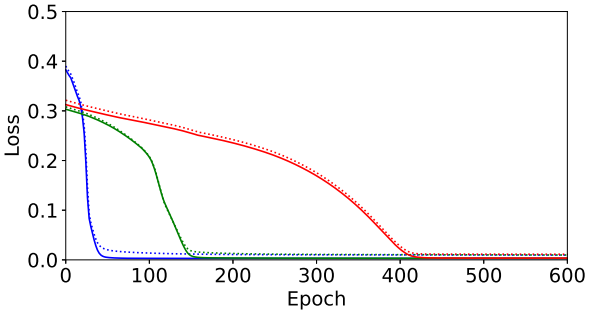
(b) Four Parameters with Six Layers



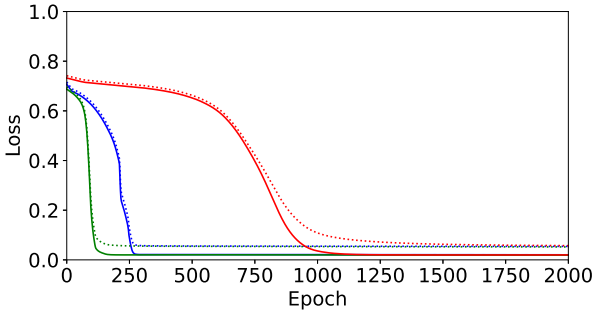
(c) Two Parameters with Eight Layers



(d) Four Parameters with Eight Layers



(e) Two Parameters with Ten Layers



(f) Four Parameters with Ten Layers

Figure 10. Training and test loss curves for two and four parameters, only depicting the first 600 and 2000 epochs, respectively, with with six, eight, and ten layers. The solid lines depict the test loss and the dashed lines depict the training loss.

We defined in Section 2.3.3 that the ideal test (and training) loss curve does not contain a sharp 'elbow' as the loss decreases; this curve should also not be linear. A smooth curve that

approaches zero is ideal [8]. One such example is in the solid red curve of learning rate of $1x10^{-7}$, in Figure 10f, which has a smooth, gradual decline at the bottom prior to flattening, implying an effective learning rate. However, the plateau-like part in the prior to the rapid decline suggests that a larger learning rate could have been applied at the start, and then switched to a smaller learning rate. This is what is referred to as a learning rate schedule [8], and could be applied in future work.

We additionally note that there are no sources of noise in the data, but that there is statistical noise from the CNN learning optimisation process. The stability of the accuracy and loss curves, however, indicates that the noise is negligible for most plots.

The peak accuracies of S_8 occurring prior to the 'elbows' of the loss curve for some models appear to be a problem with the training of the CNNs. Despite this, we can take solace in knowing that accuracy would only be inferior at later epochs, so this is a practical point at which to make the analysis.

5 Conclusions and Future Work

5.1 Future Work

As we have described, the best test loss curve has a shape between a flat line and a sharp elbow, going from high loss to low loss. We note that only a few loss curves took on this shape, and the ones that did had the lowest learning rate of the three. These plots however, began with a slow decrease for a large number of epochs before presenting the ideal curve. A good example can be found with the red line in Figure 10f. The slow decrease indicates that the initial learning rate was too small; beginning with a larger learning rate then decreasing to a smaller one could help produce an ideal loss curve, and hence improve training. We also note that the only learning rate with the correct shape loss curves was $1x10^{-7}$ and that no overfitting was suggested by these loss curves. It was noticed, however, that many of the plots had non-smooth accuracy curves, which suggests that the model was updating its weights too quickly. These observations – no overfitting and effective low learning rates – suggest that even smaller learning rates could also prove fortuitous within the structure of a learning rate schedule.

Training these CNNs for tens of thousands of epochs could yield even more precise predictions. Alternatively, with four parameter training, pre-determined weights could be given to the parameters to reflect their impact on weak lensing analysis: S_8 and Ω_m would be weighted up and h and w_0 would be weighted down. Implementation of residual neural networks – a type of artificial neural network – on the data set could yield positive results, as used in Fluri et al. 2019 on the KiDS-450 weak lensing maps [11].

We also add that future work could calculate and plot the standard deviations of the predictions. The accuracies could also be measured across all the cosmologies instead of just their combination. Future work would also benefit from focusing the testing on a larger dataset of

just one cosmology; this would allow us to learn more about individual cosmological models [41].

In this work, we have shown the effectiveness and high accuracy of our trained CNN models in predicting the S_8 cosmological parameter, but this could be taken a step further. Converting these into cosmological constraints with a cosmic shear emulator and an MCMC sampler (as discussed in Harnois-Déraps et al. 2019) would allow for predictions in between the cosmological parameter values provided in *cosmo*-SLICS [7]. The models developed in this project can only make predictions based on the cosmologies used in training. Applying the method above to this range in the parameter space could thus effectively constrain the four parameters together. Future simulations with weak lensing systematics (e.g. noise) could be developed for use in this experiment as well. Importantly, an emulator would be required in order to spot the systematics prior to the development of the CNN.

The use of modified, $f(R)$, gravity simulations provides an interesting path for further deep learning analysis. Using the already-trained best models, they could be applied to the new (2021) suite of cosmological simulations, FORGE (F-of-R Gravity Emulator) [42]. These simulations contain double the amount of nodes as *cosmo*-SLICS, meaning more data for training and testing [42]. Combining the two datasets could produce an even larger training set and very likely improve the accuracy. The results of these predictions may provide an exciting path forward.

Finally, scientists in the field look forward to gaining more data from future surveys, such as the upcoming Euclid and LSST (at the Vera C. Rubin Observatory) [7]. The data from these catalogues could be manipulated to match the map format of *cosmo*-SLICS and could be predicted on with the best developed models [7]. These are unlikely to yield positive results due to the noise in survey maps, but extensive training on future surveys with noise could prove effective. This would require CNNs that are prepared to handle noise from systematics without fitting the noise itself.

5.2 Conclusion

This work implemented novel deep learning analysis on the *cosmo*-SLICS weak lensing maps with the objective of producing high accuracy convolutional neural networks trained on a set of four cosmological parameters (S_8, Ω_m, h, and w_0). Thorough testing yielded a CNN that predicted with 97.07% confidence on the S_8 parameter when trained with the four cosmological parameters. This led to two parameter training exclusively on S_8 and Ω_m, which produced a best accuracy prediction on S_8 of 98.50%. This result reinforced the knowledge that S_8 and Ω_m have more significance in the use of weak lensing maps for cosmological parameter inference. The results of this research allowed for the proposal of a great deal of future work, suggesting that deep learning cosmic shear analysis has the potential to outperform traditional weak lensing statistical methodology used in cosmological parameter constraint.

References

[1] D H Weinberg, M J Mortonson, D J Eisenstein, C Hirata, A G Riess, and E Rozo. Observational probes of cosmic acceleration. *Physics Reports*, 530(2):87–255, sep 2013.

[2] A Riess et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009–1038, sep 1998.

[3] S Perlmutter and The Supernova Cosmology Project. Measurements of Ω and Λ from 42 high-redshift supernovae. *The Astrophysical Journal*, 517(2):565–586, jun 1999.

[4] R Mandelbaum. Weak lensing for precision cosmology. *Annual Review of Astronomy and Astrophysics*, 56(1):393–433, sep 2018.

[5] E Abdalla et al. Cosmology intertwined: A review of the particle physics, astrophysics, and cosmology associated with the cosmological tensions and anomalies. *Journal of High Energy Astrophysics*, 34:49–211, jun 2022.

[6] What is gravitational lensing? `https://sci.esa.int/web/euclid/-/what-is-gravitational-lensing-`, Accessed 24th Apr 2022.

[7] J Harnois-Déraps, B Giblin, and B Joachimi. Cosmic shear covariance matrix in wCDM: Cosmology matters. *Astronomy & Astrophysics*, 631:A160, nov 2019.

[8] L Alzubaidi, J Zhang, A J Humaidi, A Al-Dujaili, Y Duan, O Al-Shamma, J Santamaría, M A Fadhel, M Al-Amidie, and L Farhan. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8, 2021.

[9] I Goodfellow, Y Bengio, and Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

[10] J Fluri, T Kacprzak, A Lucchi, A Schneider, A Refregier, and T Hofmann. A full wCDM analysis of KiDS-1000 weak lensing maps using deep learning. *Physical Review D*, 105(8), 2022.

[11] J Fluri, T Kacprzak, A Lucchi, A Refregier, A Amara, T Hofmann, and A Schneider. Cosmological constraints with deep learning from KiDS-450 weak lensing maps. *Physical Review D*, 100(6), sep 2019.

[12] T Tröster et al. KiDS-1000 cosmology: Constraints beyond flat λCDM. *Astronomy & Astrophysics*, 649:A88, may 2021.

[13] C Heymans. KiDS-1000 cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints. *Astronomy & Astrophysics*, 646:A140, feb 2021.

[14] B Giblin, M Cataneo, B Moews, and C Heymans. On the road to percent accuracy II: Calibration of the non-linear matter power spectrum for arbitrary cosmologies. *Monthly Notices of the Royal Astronomical Society*, 490(4):4826–4840, sep 2019.

[15] J Peacock. *Cosmological physics*. Cambridge University Press, 2010.

[16] G Efstathiou. A lockdown perspective on the Hubble tension (with comments from the SH0ES team). 2020.

[17] W L Freedman. The Carnegie-Chicago Hubble program. VIII. an independent determination of the Hubble constant based on the tip of the red giant branch. *The Astrophysical Journal*, 882(1):34, aug 2019.

[18] A G Riess. A comprehensive measurement of the local value of the Hubble constant with 1 km/s/mpc uncertainty from the Hubble Space Telescope and the SH0ES team. 2021.

[19] Planck. `https://www.esa.int/Enabling_Support/Operations/Planck`, Accessed 1st May 2022.

[20] C L Bennett, D Larson, J L Weiland, and G Hinshaw. The 1% concordance Hubble constant. *The Astrophysical Journal*, 794(2):135, oct 2014.

[21] C L Bennett et al. Nine-year Wilkinson Microwave Anisotropy Probe WMAP observations: Final maps and results. *The Astrophysical Journal Supplement Series*, 208(2):20, sep 2013.

[22] T Tröster, A G Sánchez, M Asgari, C Blake, M Crocce, C Heymans, H Hildebrandt, B Joachimi, S Joudaki, A Kannawadi, C Lin, and A Wright. Cosmology from large-scale structure: Constraining Λ-CDM with BOSS. *Astronomy & Astrophysics*, 633:L10, jan 2020.

[23] T M C Abbott et al. The dark energy survey: Data release 1. *The Astrophysical Journal Supplement Series*, 239(2):18, nov 2018.

[24] C Heymans. *The Dark Universe*. 2399-2891. IOP Publishing, 2017.

[25] K L Pandey, T Karwal, and S Das. Alleviating the H_0 and σ_8 anomalies with a decaying dark matter model. *Journal of Cosmology and Astroparticle Physics*, 2020(07):026–026, jul 2020.

[26] C A P Bengaly, C Clarkson, M Kunz, and R Maartens. Null tests of the concordance model in the era of euclid and the SKA. *Physics of the Dark Universe*, 33:100856, sep 2021.

[27] H Hildebrandt et al. KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing. *Monthly Notices of the Royal Astronomical Society*, 465(2):1454–1498, nov 2016.

[28] P Simon. How accurate is Limber's equation? *Astronomy & Astrophysics*, 473(3):711–714, aug 2007.

[29] P G Castro, A F Heavens, and T D Kitching. Weak lensing analysis in three dimensions. *Phys. Rev. D*, 72:023516, Jul 2005.

[30] T Castro, M Quartin, C Giocoli, S Borgani, and K Dolag. The effect of baryons in the cosmological lensing PDFs. *Monthly Notices of the Royal Astronomical Society*, 478(1):1305–1325, may 2018.

[31] H Hildebrandt et al. KiDS-1000 catalogue: Redshift distributions and their calibration. *Astronomy & Astrophysics*, 647:A124, mar 2021.

[32] M Bartelmann and P Schneider. Weak gravitational lensing. *Physics Reports*, 340(4-5):291–472, jan 2001.

[33] J Harnois-Déraps, N Martinet, T Castro, K Dolag, B Giblin, C Heymans, H Hildebrandt, and Q Xia. Cosmic shear cosmology beyond two-point statistics: a combined peak count and correlation function analysis of DES-y1. *Monthly Notices of the Royal Astronomical Society*, 506(2):1623–1650, jun 2021.

[34] S Heydenreich, B Brück, and J Harnois-Déraps. Persistent homology in cosmic shear: Constraining parameters with topological data analysis. *Astronomy & Astrophysics*, 648:A74, apr 2021.

[35] C T Davies, M Cautun, B Giblin, B Li, J Harnois-Déraps, and Y Cai. Constraining cosmology with weak lensing voids. *Monthly Notices of the Royal Astronomical Society*, aug 2021.

[36] Christopher T. Davies, Marius Cautun, Benjamin Giblin, Baojiu Li, Joachim Harnois-Déraps, and Yan-Chuan Cai. Cosmological forecasts with the clustering of weak lensing peaks. 2021.

[37] A Paszke et al. Pytorch: An imperative style, high-performance deep learning library. In H Wallach, H Larochelle, A Beygelzimer, F d'Alché-Buc, E Fox, and R Garnett, editors, *Advances in Neural Information Processing Systems 32*, pages 8024–8035. Curran Associates, Inc., 2019.

[38] G Van Rossum and F L Drake. *Python 3 Reference Manual*. CreateSpace, Scotts Valley, CA, 2009.

[39] The pandas development team. pandas-dev/pandas: Pandas, February 2020. `https://doi.org/10.5281/zenodo.3509134`, Accessed 21st Feb 2022.

[40] S Mostafa, D Mondal, M Beck, C Bidinosti, C Henry, and I Stavness. Visualizing feature maps for model selection in convolutional neural networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, pages 1362–1371, oct 2021.

[41] J Harnois-Dérap et al. Cosmological simulations for combined-probe analyses: covariance and neighbour-exclusion bias. *Monthly Notices of the Royal Astronomical Society*, 481(1):1337–1367, aug 2018.

[42] C Arnold, B Li, B Giblin, J Harnois-Déraps, and Y Cai. Forge – the f(r) gravity cosmic emulator project i: Introduction and matter power spectrum emulator. 2021.