Tilburg University

Master's Thesis

MSc Economics

Track: Data Science

Can Satellite Images and Neural Networks be Used to Estimate Industrial Production?

July 2020

Author

Thomas Chatt

Student number

****

Supervisor

Dr. Madina Kurmangaliyeva

Second reader

Dr. Marie Le Mouel

Abstract

Due to the time lag and inaccuracy of official GDP statistics a methodology of using the change in satellite Luminosity data is a relatively new way to complement official statistics. This thesis builds on previous literature by creating a methodology which uses Machine Learning instead of traditional econometrics. Machine learning and specifically Convolutional Neural Networks are implemented to estimate how changes in Luminosity can predict changes in Industrial Production. Industrial Production is used as a proxy for GDP due to being published at monthly not quarterly intervals giving more observations to train the model. This methodology is then applied to estimate the fall in the Industrial Production in China, due to the COVID-19 pandemic. The results are promising, with the model yielding low losses (the difference between estimates and true value) and accurate predictions. Finally, this thesis finishes by presenting a range of possible improvements which could be incorporated to potentially yield more accurate predictions and contribute to a full investigation regarding the effectiveness of Neural Networks for predicting these output related outcome variables.

Acknowledgements

I would like to take this opportunity to express my sincere thanks and appreciation to those who have guided, supported and motivated me throughout both this thesis and academic year. In particular, I would like to thank my thesis supervisor Dr. Madina Kurmangaliyeva for her constant valuable suggestions, feedback and guidance throughout the thesis process.

I would also like to thank my friends and family and specifically my parents for their support during not only the thesis period but throughout the last four years of University. I would not have been able to do it without them. Furthermore, I would like to extend my gratitude for the community of Stack Overflow, who helped me to overcome so many of the coding challenges which I encountered during the last few months.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Using changes in satellite imagery to predict changes in output is a relatively recent development, with a number of the most relevant papers being covered in section 2.1. The aim of these papers is to provide a more accurate measurement of GDP (Gross Domestic Product) growth in countries where the national statistics are not reputable. These previous papers take simple averages of very large Luminosity data-sets and use this single value to run in OLS and fixed effects panel regressions.

The objective of this thesis is to apply Luminosity data from satellites and make predictions using a Deep Neural Network (DNN) instead of econometrics. Using machine learning and a DNN will enable the model to take advantage of the information available within the full satellite images. Therefore, these DNN's should be able to extract more information and thus produce more accurate predictions than those generated from traditional econometrics.

More specifically, this thesis will apply this methodology to the case of China in quarter 1 of 2020. At this time, China was experiencing it's peak of the COVID-19 pandemic, which has subsequently dominated both the news and people's consciences, with this unlikely to change in the coming months or years (Campbell, 2020).

China is a very interesting case since it has the largest population, second largest GDP (World Bank, 2018b), and is the largest global exporter (World Bank, 2018a).

Therefore, accurately analysing the economic effects of the pandemic in China will have large positive spillovers for the global economy. Other countries will be able to use this methodology to more accurately predict the effects of reduced production and thus reduced exports from China, on their own economy. In addition to these secondary effects, these countries will also be able to better predict the likely impact of reduced production within their own economy.

Furthermore, using satellite images means that China (or any other country or organisation who apply this methodology) will be able to assess the economic damage they have endured much sooner than if they were to wait for official GDP statistics. Output is conventionally measured using GDP, but due to the large time and resources needed to estimate this, it is published with a time lag and often only once a quarter. With the methodology introduced in this thesis, the satellite data is available a number of days after the recording date. Therefore, pre-training a model on past data with past GDP statistics will allow analysts to predict the contractions in output with only a short time lag.

In addition, countries which are comparatively recently experiencing their spikes in COVID-19 cases will be able to use satellite images from a range of countries, to compare the speed at which their recoveries took place. This may mean that governments will be able to better choose which policy responses they undertake, based upon past successes in other countries.

The economic impacts of COVID-19 are unlikely to be the same as recent crises borne from within the economic system. COVID-19 is exogenous from this economic system and therefore the reaction and precautions taken by governments have been very different from the 2007/2008 financial crisis. In addition, due to the scale of this crisis, it may be the case that an elasticity estimated with traditional econometrics is not flexible enough to accurately be applied in these circumstances. Therefore, using a machine learning methodology may improve the ability of analysts to accurately predict changes in output in these more extreme cases.

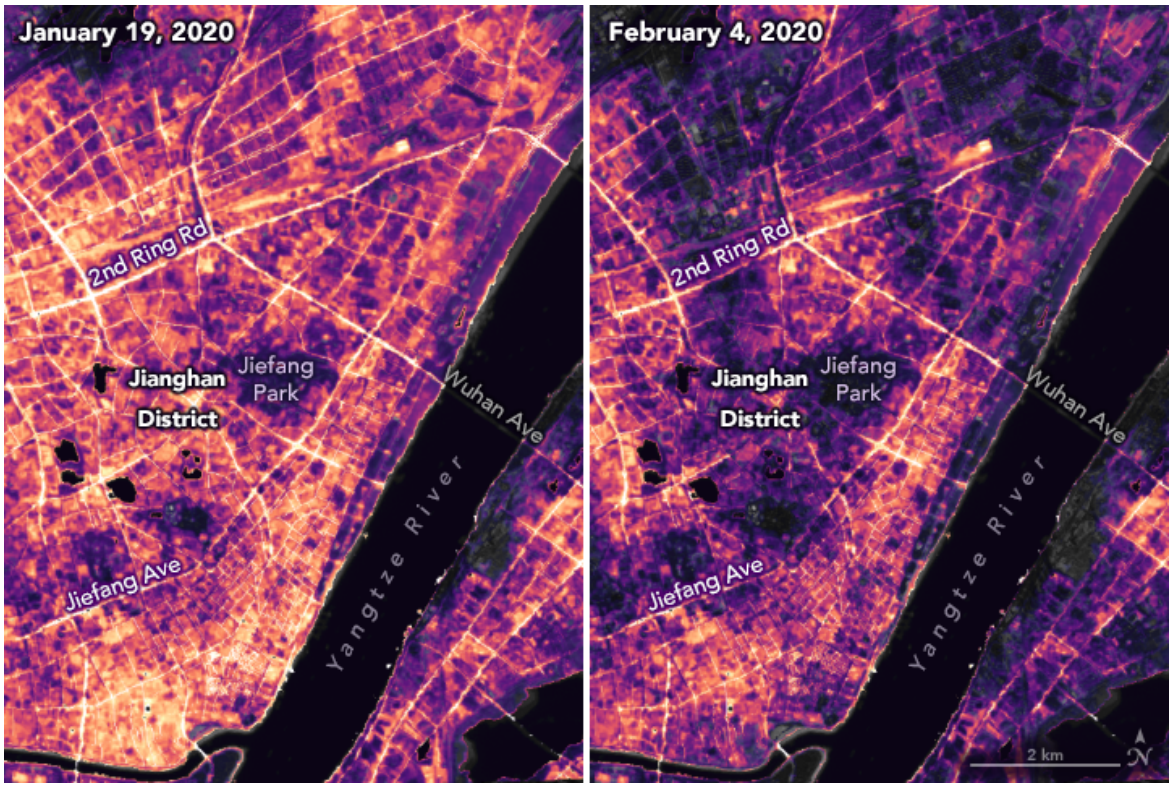


Figure 1.1: Jianghan District January vs February 2020 (Carlowicz, 2020)

As reported by Department of Health (2020), the first cases of COVID-19 were in Wuhan in Hubei Province, China. Figure 1.1 is an example of the information which can be gleaned from satellite Luminosity images. This image is a representation of the night lights in Wuhan at a very high resolution, one which is not available to the public. Figure 1.1 shows the large decrease in Luminosity which was experienced in Wuhan between January and February 2020. The significant decrease is especially prevalent in the areas of Jiangham District (the central district) as well the areas around 2nd Ring Road. This decrease in Luminosity is likely to be associated with the significant and severe fall in economic activity due to the lockdown.

Within the following sections of this thesis, a methodology for predicting changes in output from changes in satellite Luminosity will be developed. It will describe a range of issues which have been encountered and considered, as well as the results from a number of machine learning models. In addition, in section 5.1 there are a number of suggestions as to how this methodology could be further improved upon and tested in the future, to yield increased accuracy and gain a more complete picture into its effectiveness.

Chapter 2

Literature Review

2.1 Literature Review

There is extensive literature published on the topic of using satellite images to predict economic outcomes. Some of the most relevant papers to this field and research question include Henderson et al. (2012), Bickenbach et al. (2016), Chen and Nordhaus (2011) and Jean et al. (2016).

Henderson et al. (2012) develop a framework which uses satellite data of night lights to improve the accuracy of officially published income growth statistics. They use a sample of over 100 low and middle income countries to show that there is a stable and statistically significant relationship between the growth rate of luminosity (measuring night lights) and the growth rate of GDP. Using satellite data also provides the possibility of enabling measurements to be conducted at regional and even sub regional levels. This is a weakness encountered with traditional reporting of GDP, which is measured at national level. The data used in Henderson et al. (2012) was collected from the National Oceanic and Atmospheric Administration's National Geophysical Data Center (NOAA-NGDC). NOAA removed observations for places in the bright half of the lunar cycle (or in other words in summer), areas experiencing Aurora activity (northern lights) and forest fires. The luminosity data is represented by pixel as a number between 0 and 63, where a pixel reporting zero shows no luminosity and 63 shows the highest luminosity.

Henderson et al. (2012) also note that measured luminosity for the same GDP can vary with changes in the composition of production and the division of economic activity between the day and night, as well as population density.

The formalisation of this best fit elasticity has been included in full in section 2.1.1 since this methodology is used in baseline comparisons in this thesis. Equations are also included which allow for year fixed effects and cross country cultural effects for differences in light usage [1]. However, this is not important for this thesis since it only analyses China and as such, cross country controls are not necessary.

Henderson et al. (2012) find that their estimates differ from official reported GDP growth by up to three percentage points, as well as finding a point estimate of the elasticity of night light growth with respect to GDP growth of 0.3. Since this elasticity is less than 1, the percentage increase in luminosity is larger than the percentage increase in income. They also find that lights respond asymmetrically to changes in incomes meaning that luminosity increases more in times of economic growth than it falls in periods of downturn.

However, the methodology used in this paper is not without fault. The satellites cannot differentiate between private lights (lights from cars and homes) and public lights (street lights etc.). A further technical caveat with this methodology used in Henderson et al., is the assumption that, for the growth rate reported in national accounts, the measurement error is uncorrelated with the measurement error when the change in luminosity is the explanatory variable. These errors could be correlated for a number of reasons, if there was an increase in undeclared output/income (potentially for the purposes of tax avoidance). The increase in this undisclosed income may be used to buy electronic appliances which would increase the error term of luminosity growth on income growth since there would be an increase in luminosity caused by these appliances and also would not be reported in official statistics for GDP growth.

The methodology also runs into technical issues with the need to control for cross-country cultural differences relating to light usage. However, since this thesis is only

[1] This can intuitively be thought of as different cultures having different approaches to light uses. In the west, it is common for households to use lights in their homes when it is dark, whereas in rural Africa, oil lamps and candles are more commonly seen.

examining the relationship in one country (China) this will not be an issue. Furthermore, measuring GDP using luminosity may not be an accurate measure of true economic activity, since changes in economic activity at night (proxied by luminosity) may not be representative of activity in the daytime.

Bickenbach et al. (2016) expand on the work of Henderson et al. (2012) by seeking to assess if there is a relationship between growth in luminosity and growth of GDP at a regional and sub-regional level. In order to try and ascertain if this relationship holds, they examine data from Brazil and India and break it down into 5 regions per country. They estimate the main equation (long term relationship between night lights growth and true GDP growth) from Henderson et al. (2012) equation (6) (formalised in section 2.1.1) but found differing results. The regional estimates of β_0 may be subject to bias from systematic differences in measurement error of GDP between regions.

In order to try and overcome these issues they extend the model by controlling for region specific changes by adding in dummy variables for all but one of the 5 regions. This is formalised by $D_r, r = 2, ..., R$. They also extend the baseline model by controlling for measurement errors across space, by adding a spatial lag for light growth. It is hypothesised that the error term in equation (6) (in section 2.1.1), follows the form: $u_i = \gamma_i l_i + u_{i0}$ where u_{i0} has a zero expected value and country specific variances, γ_i, are correlated across countries. Bickenbach et al. (2016) also hypothesise that these errors are more similar in countries located geographically nearby to country i. Incorporating regional dummies and spatial lags captures effects of determinants of the structural growth of country i which may have been omitted. This was also proposed in Berliant and Weiss (2017).

Berliant and Weiss find that the parameter estimate for β_0 varies significantly between regions, with estimates as low as 0.1(significantly smaller than 0.3 found by Henderson et al.). The paper also reports that India has large variation in estimates of the parameter β_0 between different regions and these are not stable over time. The authors also find that there is not a stable relationship between night light growth and real GDP growth between regions in both Brazil or India. This is one of the main differences between this paper and the findings of Henderson et al. (2012). They do not find a stable relationship between the growth of night light intensity and GDP growth at

regional and sub regional levels. They also find that this is not caused by bias from measurement in GDP, as this was controlled for separately with the same result.

Chen and Nordhaus (2011) examine how poor quality data has caused issues in understanding economic growth and whether luminosity data increases accuracy at either regional or sub-regional levels. The paper also uses data from the Defense Meteorological Satellite Program Operational Linescale System (DMSP-OLS). They looked at two types of data from this database:

- Stable - removes short lasting events such as fires, as well as background noise.
- Calibrated - cleaned data but released much later (this paper was published in 2011 and the latest calibrated data was only from 2006).

They use sub-regional data from the G-Econ data set provided by Yale University, to use as a baseline for the output in local areas. G-Econ has data on gross output for grid cells which are 1° latitude by 1° longitude (111KM x 111KM in size).

The analysis begins in a similar way to both Henderson et al. (2012) and Bickenbach et al. (2016), the equations are:

- (1) $y_i = y_i^* + \epsilon_i$ - the error process for output.
- (2) $m_i = m_i^* + \eta_i$ - the error process for luminosity.
- (3) $m_i = \beta y_i^2 + u_i$ - the data generating process for luminosity as a function of output.

The methodology used in Chen and Nordhaus then differs from the aforementioned papers. They invert equation (3) which becomes $\hat{z}_i = (\frac{1}{\hat{\beta}})m_i$ where $\hat{z}_i$ is the log of the luminosity-output proxy and $\hat{\beta}$ is the estimated coefficient in (3).

They break their analysis down into cross-sectional and time series methodologies. They find that the error estimates are consistent with the Penn World Tables (PWT), where countries ranking lower on the PWT tables have higher error estimates of reported GDP. Luminosity adds a large amount of information for countries rated D, which include Algeria, Cambodia and the Democratic Republic of Congo. E countries

(the worst ranked) include Myanmar and North Korea.

The authors also estimate a short-term GDP-light growth relationship using a panel fixed effect approach and annual data. The estimated equation is $lnY_{it} = \alpha + \beta_0 ln(L_{it}) + \delta_i + \delta_t + u_{it}$, which is the standard equation (6) with the addition of country fixed effects δ_i and year fixed effects δ_t. However, they did not find the elasticity to be constant over time and between countries.

Dai et al. (2017) use both the DMSP-OLS satellite data as well the global Suomi National Polar-orbiting Partnership (NPP) night-time light data. The NPP data is from a more modern satellite and as such has benefits including a higher spatial resolution and a wider radio magnet detection range. The authors compare the suitability of the two data sets when estimating the relationships in mainland China. They use the DMSP/OLS stable light data which has been cleaned and is cloud free. The GDP data for provincial and city level areas is sourced from the China City Statistical Yearbook of 2014 and the Local Statistical Yearbook. The authors use three different estimation equations (linear regression, power function and polynomial model). The equations are formalised below:

- $\hat{GDP}_i = a \cdot TNL + b$ - linear
- $\hat{GDP}_i = a \cdot TNL^b$ - power function
- $\hat{GDP}_i = a \cdot TNL^2 + b \cdot TNL + c$ - polynomial model

Where TNL represents the Total Nighttime Light, which is the sum of all pixel values in each administrative unit. The authors find that the NPP/VIIRS data set provides higher correlation coefficients and lower residuals for national and provincial aggregation levels. However, for city levels this does not hold. This does not pose an issue since this thesis will examine the relationship at a national level. In addition, Bickenbach et al. (2016) show that the relationship between luminosity growth and GDP growth does not hold in regional and sub regional agglomerations anyway. The higher resolution of the NPP/VIIRS data leads to more accurate GDP prediction than the typical DMSP/OLS data due to its higher radiation resolution.

Liu (2006) finds a unidirectional Granger causal relationship between ln(NO_x) (ni-

trogen oxides) and ln(GDP). Granger causality seeks to assess if there is causality between two time series by looking at a cause and an effect. If it is hypothesised that one variable (X) causes changes in (Y) then changes in X should precede changes in Y. Therefore, NO_x data could be included in estimations of GDP.

Applying machine learning to predict economic outcomes has also been used with some success in the paper Jean et al. (2016). They apply machine learning to predict poverty in a range of African countries. Overall, even with inexact data, Jean et al. find what they describe as 'quite accurate' predictions for poverty in these countries. Furthermore, interestingly, they find that the predictive power of their models declines only by a small amount when a model trained on data from one country is used to estimate consumption or assets in other countries.

2.1.1 Formalisation of Equations From Past Literature To Be Used In This Thesis

In section 3.4, methodologies from previous literature have been used in this thesis to compute a baseline OLS (Ordinary Least Squares) estimate of the elasticity of luminosity growth to GDP growth. The equations and assumptions from past literature are formalised below.

Henderson et al. (2012) formalise their best fit elasticity in the following way:

- Measurement error in GDP growth reported in national accounts is expressed in the following equation: $z_j = y_j + \epsilon_{z,j}$ (4) where z_j is the growth of real GDP reported in country $j's$ national accounts, y_j is the growth in true real GDP and $\epsilon_{z,j}$ is the error term in this reported statistic.

- The relationship between growth of lights and growth of true income is given by $x_j = \beta y_j + \epsilon_{x,j}$ (5), where x_j is the growth in observed light.

  - Where ϵ_x is denoted σ_x^2

- It is assumed that there is a constant elasticity relationship between total observable lights (X) and total income (Y), formalised by $X_j = Y_j^{\beta}$ and β is the estimated elasticity of lights on income.

Henderson et al. test different functional forms with a range of controls for changes in the dispersion of light and conclude that eq (5) is an appropriate estimation equation.

The error term in equation (5) encompasses different measurement errors. These include:

- Errors in measuring Luminosity, the difference between the true emanating light from earth into space and the light being measured by the satellite.
- Variations between countries in the core relationship between the growth of GDP and the growth of luminosity. This relationship variation can be attributed to differences in the main sectors which are expanding.

The regression for predictive purposes (predicting the growth of income on the growth of lights) is $z_i = \hat{\psi} x_j + e_j$:

- The parameter $\hat{\psi} = \frac{cov(x,z)}{var(x)}$ which is the estimate of the inverse elasticity of luminosity with respect to income and is estimated using OLS.
- The relationship between $\hat{\psi}$ and the structural parameter β has $plim(\hat{\psi}) = \frac{1}{\beta}(\frac{\beta^2\sigma_y^2}{\beta^2\sigma_y^2+\sigma_x^2})$ which is the inverse of the elasticity of lights with respect to income. σ_y^2 is the variation of true income growth and σ_x^2 is variation in luminosity growth.

The long term relationship between growth of luminosity and GDP growth can be formalised in the following way (where each country is represented by i) (formalisation taken from Bickenbach et al. (2016)):

- $y_i^* = \beta_0 l_i + u_i$ (6) where $u_i = \alpha + \epsilon_i$ where y_i^* is the observable true growth rate of country i for a given time period. l_i is the contemporaneous growth rate of night light intensity. β_0 becomes the parameter of interest. α represents a national growth component in the error term.

Chapter 3

Methodology and Data

3.1 Methodology and Data

This thesis differs from previous literature and methodology, since instead of using traditional econometric analysis it will explore how and whether Artificial Neural Networks (ANN)/Deep Learning and Big Data [1] can improve on the previous methodologies predicting economic output.

3.2 Luminosity Data

The Luminosity data used for this thesis has been collected from the Visible Infrared Imaging Radiometer Suite (VIIRS) probe on board the Suomi National Polar-orbiting Partnership (NPP) satellite (Román, 2020). From the official release document (Román et al., 2019), the unit of measurement for Luminosity in this satellite is nano-watt per steradian (angle to the centre of the sphere) per square centimetre $nW \cdot cm^{-2} \cdot sr^{-1}$. The NPP satellite was chosen instead of DMSP-OLS due to the findings from Dai et al. (2017). They found that the smaller size of the pixels and increased range of Luminosity values (from 0 to 65,534, vs 0 to 63) both lead to enhanced detail in the data. Therefore, these improvements should lead to better predictions for the outcome variable, unless there is more noise in the data.

The product code is VNP46A1, which contains a range of data-sets in the visible

[1] Big Data refers to unstructured data such as images, films and audio.

and near infrared light spectrum. The data-set from VNP46A1 used in this thesis is 'DNB At Sensor Radiance 500m' where DNB stands for 'Day Night Band'. This data-set is unprocessed and therefore includes clouds and other natural phenomena such as Aurora. It would have been preferable to use VNP46A2 data, which has been processed to remove these phenomena from the data set. However, this has not been released to the public yet. A section of VNP46A2 data recorded in 2012 will be released in summer 2020. Figure 3.1 shows the effectiveness of the processing used in VNP46A2.

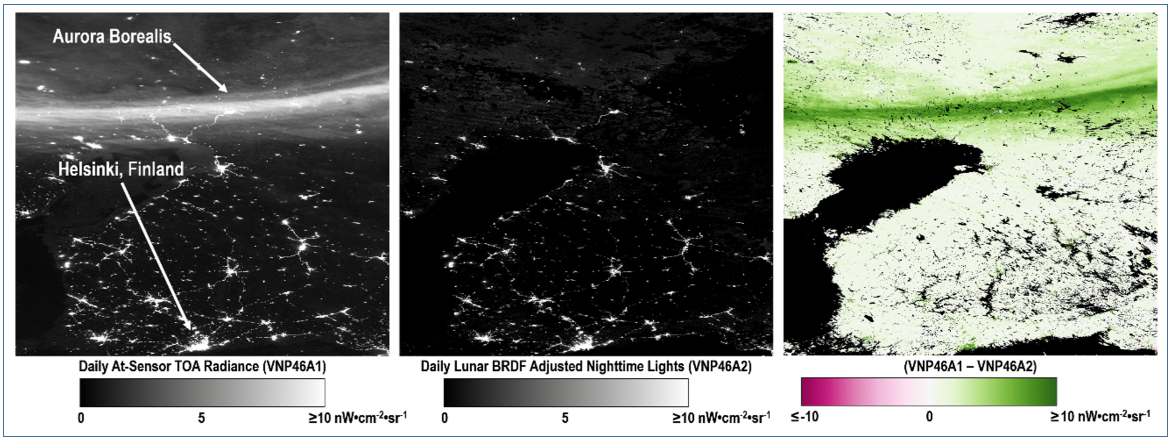


Figure 3.1: VNP46A1 vs VNP46A2, from NASA (2020)

It is only possible to download daily data from the NASA website and in the case of VNP46A1, it is split up into blocks of 1200KM x 1200KM with pixels of size 500m x 500m.

Figures 3.2 (as well as figures A.1 and A.2 from appendix A) have been produced by concatenating the blocks in the geographically correct order using their individual block references (which correspond to their geographical locations on Earth [2]) to create an image of the the night lights (Luminosity) in China on individual days. A representation of this is as follows:

[2]The order of these blocks is shown in appendix A in table A.1

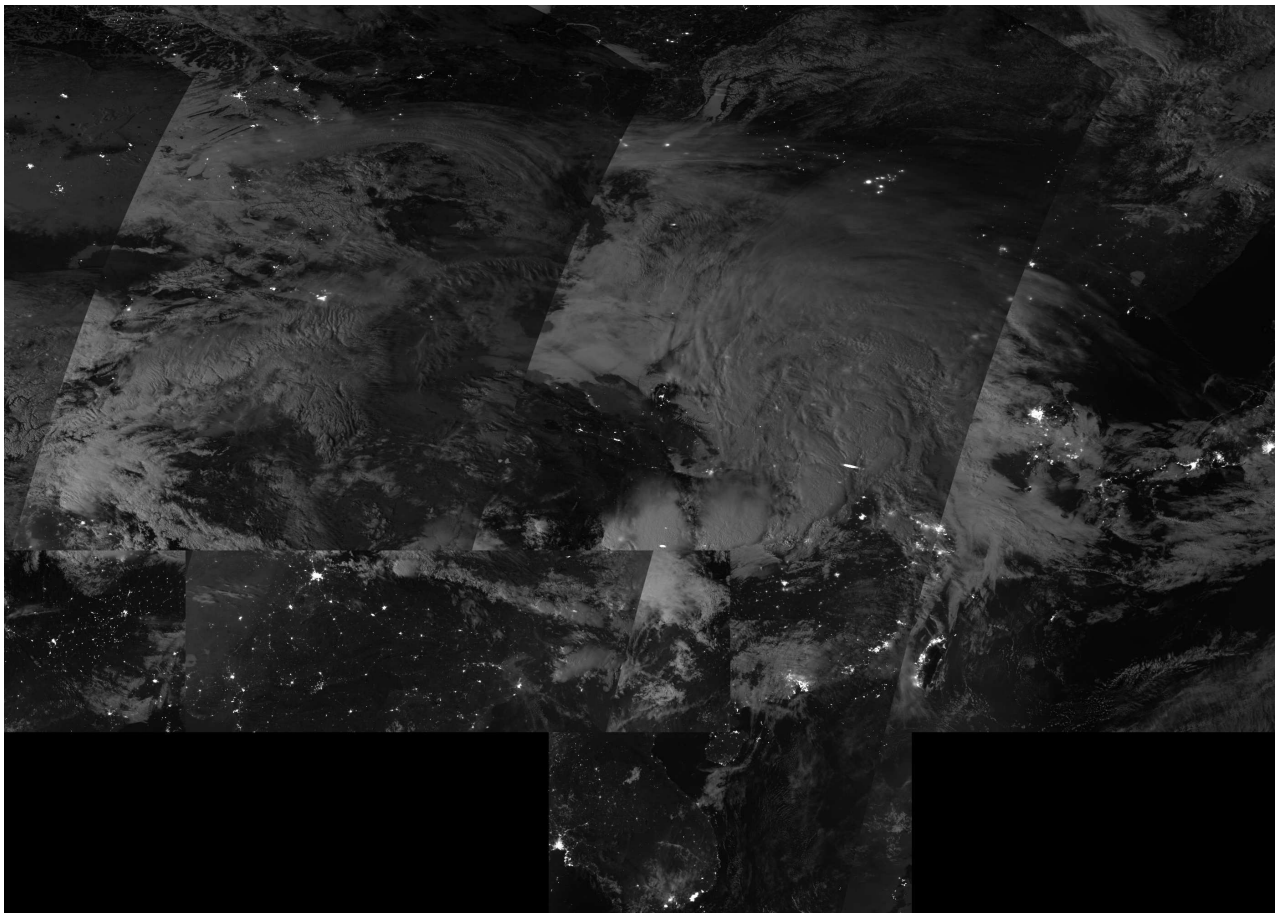


Figure 3.2: 1^{st} April 2020 Night Lights over China

Figures 3.2, A.1 and A.2 have been selected since they show how noise and specifically in this case cloud cover, can affect the images which are used to create the predictions. Figure 3.2 has moderate cloud cover over large areas of China. Figure A.1 (from appendix A) is an optimal image with no discernible cloud cover. This image clearly shows the areas with the highest irradiating night lights.

An example of an image which illustrates a high noise day is figure A.2 (from appendix A). This image is distorted by very bright clouds, it is not possible to see some of the largest settlements and economic hubs located on the East coast of China since they are being covered by this cloud. In addition, the overall image appears to be brighter, even for areas which are not covered by cloud. This noise is likely to have a negative impact on the results as the ANN will think of these bright areas as areas with high Industrial Production.

Since the Industrial Production data is published at monthly intervals, it is necessary

to average the VNP46A1 blocks to become monthly averages. The monthly averages for February 2019 and 2020 have been concatenated (joined together) to present the images with their blocks in the correct geographical locations. Using this averaged data reduces the noise which is present on days with large amounts of cloud cover and therefore should lead to more accurate predictions.

In order for the Convolutional Neural Network (CNN) (explained in section 3.5) to be able to learn from these independent blocks, they must be concatenated to create one wide image, with each full month block containing individual blocks in the same order. This enables the CNN to learn how the data in each block changes with a change in the outcome variable and therefore should lead to a more accurate model. An example of one month's averaged and concatenated data presented in figure A.3 in appendix A.

It is interesting to compare the distribution of Luminosity values between February 2019 and February 2020, since it may provide insight into changes in light usage due to the COVID-19 pandemic.

Using the full range of Luminosity values (0-65,534) in a histogram did not provide any useful insight, even with a large number of bins [3]. Due to this, the ranges 0-1500 and 0-100 have been chosen for the plots on the histograms below. In February 2019, only 8,531 of 172,800,000 (0.000049%) values are outside the range 0-1500 and only 1,166,338 of 172,800,000 (0.675%) are outside the range 0-100.

[3]Bins in a histogram refer to the groups of equal width that the values (in this case for Luminosity) fall into. It was found that using more bins splits these data points across more groups which helps to show the difference in the density of the data between the two years. This is especially true in this data where the large majority of values fall in a very small range (0-100).

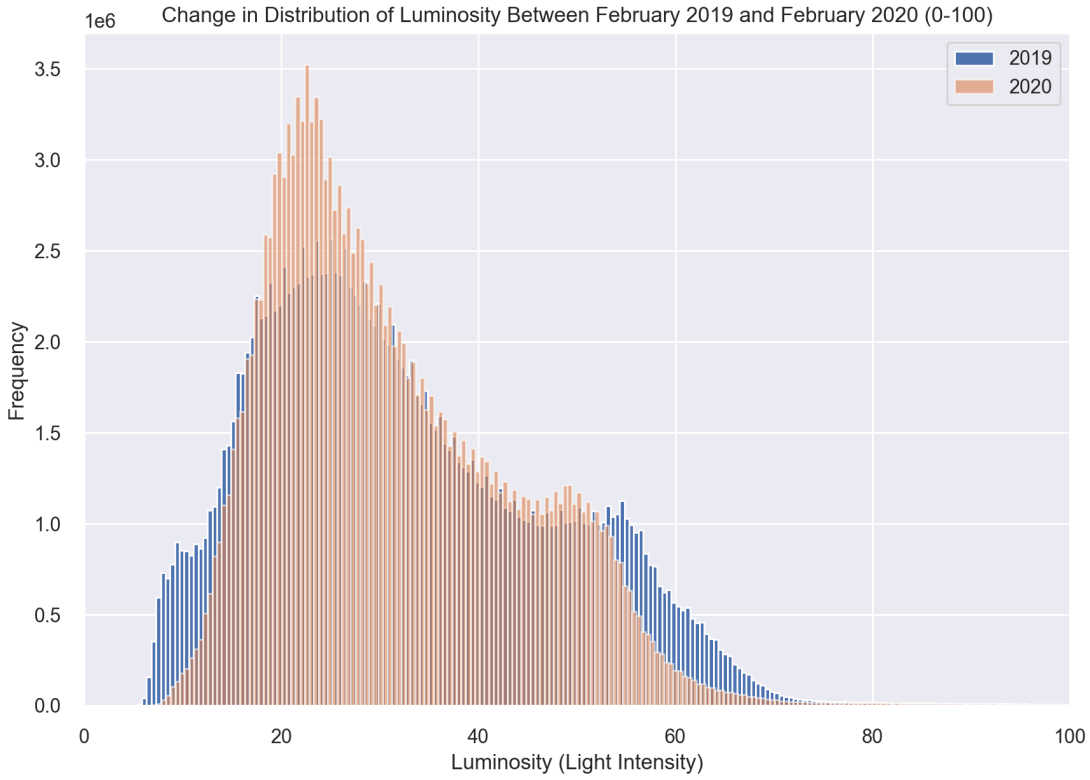


Figure 3.3: Histogram Comparing February 2019 and February 2020 Luminosity Data (Range 0-100).

Figure 3.3 (see figure B.1 in appendix B for the range of Luminosity values 0-1500) is a histograms which shows the difference in the distribution of the Luminosity values in February 2019 and February 2020 (referred to as 2019 and 2020 in the remainder of this section). For both years, the distributions have a heavy positive skew. The distribution in 2019 is more broad based, meaning that the values are spread across a larger range. There are more values in the range 8-18 and 52-70 in 2019 than in 2020 and in 2020, there is a higher peak frequency (the number of observations per bin). The difference in the maximum number of values per bin between 2019 and 2020 is approximately 1.2 million pixels.

This shows that there has been an increase in Luminosity in areas with previously very low values, and a decrease in Luminosity in areas with previously higher than average values. Since increased Luminosity suggests increased income, the normalisation of Luminosity values could potentially imply an increase in equality of income between

regions. On the other hand, the change in distribution of values may be the result of industrial areas being shut down and no longer emitting light pollution. This reduction in output also coincides with Chinese New Year, which could increase Luminosity values in rural areas due to city workers returning to celebrate with families. Figure 3.3 could also suggest that due to COVID-19, areas with high population density (and therefore high Luminosity) have reduced how much light they emit due to there being less economic activity caused by self isolation bought about by the virus.

3.2.1 Reducing data size

The original data-set was too large to run on a normal PC and there were many complications with using cloud computing services. Google's Cloud computing suite ironically made it difficult to combine with the data stored in Google Drive as well as being expensive (approximately $17 per hour). Therefore, it was necessary to reduce the size of the data-set. One methodology to do this was to take the average of a sub-matrix of pixels and report back one pixel, therefore reducing the size of the data-set (explained in figures 3.4 and 3.5). The first reduction used took a 5x5 group of pixels and averaged to one pixel. This reduced each monthly average Luminosity data-set (shown in figure A.3) to 480x14400 pixels. This approach is very similar to that of MaxPooling (but instead of taking the max value, it takes the average value), explained in detail in section 3.5 and represented in figure 3.12. However, it was still too large to run on Google Colab or a PC with 32GB of RAM.

Therefore, the next logical step was to reduce the size of the monthly data even further. A reduction of 10x10 group of pixels to one pixel was used next, where each pixel now represents an area of 5KM^2. This reduces the dimensions of each monthly data-set to 240x7200 pixels and allows the neural network to run. A graphical representation of this is below:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots \\ \vdots & \ddots & \\ a_{K1} & & a_{KK} \end{bmatrix} \longrightarrow [\bar{A}]$$

Figure 3.4: Reducing Dimensions of Array Using The Average Value

Where A represents a sub section of the full data and $\bar{A}$ represents the averaged single

pixel from array A. In this example $K = 10$ since every 10x10 block is averaged to one value.

Potentially, more accurate predictions may be produced by reducing pixels by outputting the maximum value in the block (the same as MaxPooling). In this case, for each 10x10 block the reported pixel in the reduced array would be the maximum value from the original block (see figure 3.12):

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots \\ \vdots & \ddots & \\ a_{K1} & & a_{KK} \end{bmatrix} \longrightarrow \max[A]$$

Figure 3.5: Reducing Dimensions of Array Using The Maximum Value

In order to ensure that this data was still similar to the original data and to make sure that detail and fidelity had not been lost, further descriptive statistics were run. Figure 3.6 (and table B.1 in appendix B) show the comparison of the full data-set and both the reduced to 10% data-set's (1 reduced using the max value and the other using the average value) for Feb 2019 and Feb 2020.

Comparing the mean averaged data vs the full data:
Although the mean, 25%, 50% and 75% values are very similar, the minimum and maximum values are very different by construction, due to the averaging process. The top plot in figure 3.6 represents this clearly. The range of values has decreased significantly, shown by the much lower peaks in the top box plot. This implies that the fidelity of the data has decreased and the CNN may not be able to gain as much information as to how changes in Luminosity affect changes in the Industrial Production in the brightest areas of China.

Comparing the max data vs the full data:
As expected, the max data-set retains more of the range when compared to the averaged data. However, the mean, min, 25%, 50% and 75% are very different from the full and averaged data-set as a result of the averaging process.

A number of approaches have been considered for importing the data into the CNN. The data that will provide the best results will have the largest number of unique values while keeping the structure of the data as similar to the full data-set.

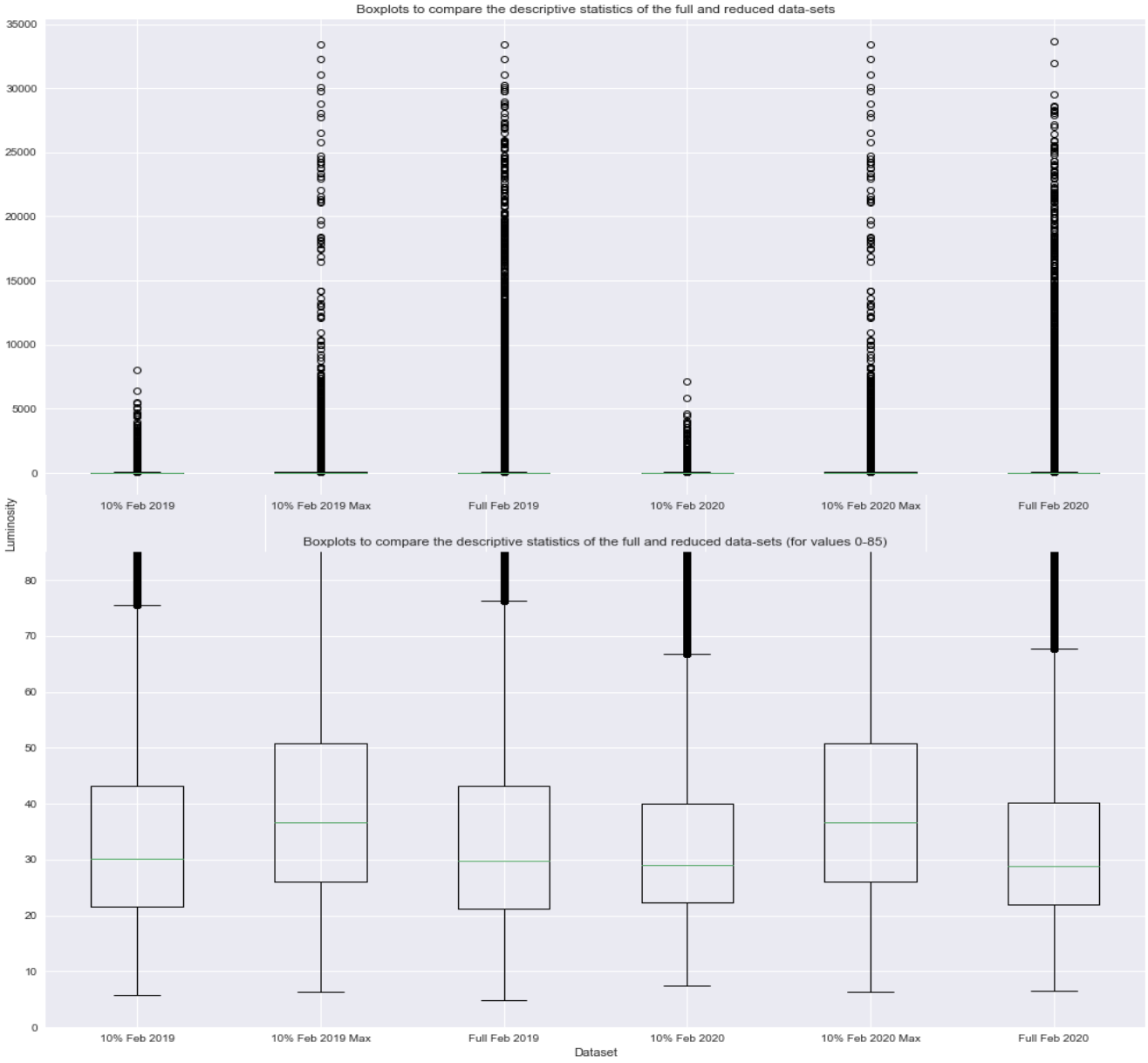


Figure 3.6: Box Plot Comparing Full vs Reduced Data

A common way to import data into a CNN is by importing them as images. However, in the context of this data, with such a large range of Luminosity values, a substantial amount of the variation in the data is lost. Table B.2 in appendix B contains information relating to methodologies trialled in this thesis to reduce the dimensions of the data while minimising the loss of information for the NN.

Normalising the values to be in the range of 0-255 is typical practice for grey scale

images. [4]

3.3 Industrial Production Data

Henderson et al. (2012) use GDP as their outcome variable. However, unfortunately this is only reported quarterly. VNP46A1 is only available from 2012 and since GDP is reported at a quarterly interval, this would only provide 32 observations (from Q1 2012 to Q4 2019), which would have to be split into training and testing sections for the Neural Network (explained further in section 3.5). In addition, VNP46A1 data is very large in size and even 5 years of data (from January 2015 to April 2020 as used in this thesis) totalled over 2TB of raw unprocessed data and led to processing power issues when running the Neural Network. Increasing this data-set to include 2012-2014 would require even more computing power to process and therefore, it is not realistically achievable for this thesis. As such, Industrial Production data will be used as a proxy for GDP. Industrial Production is defined as 'the output of industrial establishments and covers sectors such as mining, manufacturing, electricity, gas and steam and air-conditioning. This indicator is measured in an index based on a reference period that expresses change in the volume of production output' (OECD, 2020).

Industrial Production as a percentage of GDP is comparatively high in the case of China. Globally, according to the World Bank indicator 'Industry (including construction), value added (% of GDP)' (World Bank, 2018c), Industrial Production as a percentage of GDP in 2018 was 27.81% but in China it was 39.687% . In the case of China, there has been a decrease in the Industrial Production to GDP ratio since 2006 where it peaked at 47.557%. Therefore, it is likely that in the future, using Industrial Production as a proxy for GDP will become less representative of economic activity as a whole.

Industrial Production is released at monthly intervals and therefore for the time period 2015-2019 there are 60 observations. This data was retrieved from World Bank (2020), the World Bank's Global Economic Monitor data set which contains a range of Economic indicators. Industrial Production data for China from the World Bank has only

[4]Unlike colour images which contain three channels (Red, Green and Blue), Gray scale images only have one channel and produce pixels in the range of black (0) and white (255)

been published up to July 2019, giving a total of 55 observations in this data-set.

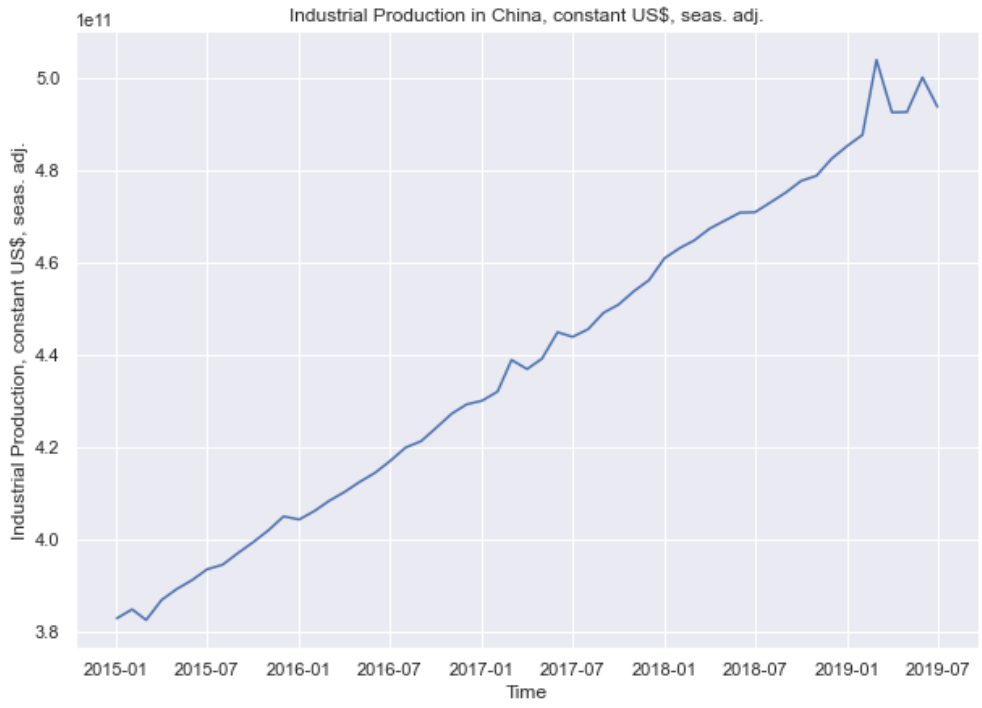


Figure 3.7: China Industrial Production, Jan 2015 - July 2019

Figure 3.7 shows the largely consistent increase in Industrial Production in China during the second half of the last decade. However, as shown from the figure, Industrial Production decreased compared to the previous period six times for reasons which are not explicitly clear. The largest decrease is in 2019 where the overall trend seems to slow down and become less stable. This could potentially be linked with China's trade war with the USA (BBC News, 2020). In just a four and a half year period Industrial Production in China increased by 28.94%, a substantial amount considering that this does not include the increase in the services sector or changes to the primary sector (which includes farming and mining).

The Chinese National Bureau of Statistics (NBS) also publishes Industrial Production with an indicator called 'Value-added of Industry, Accumulated Growth Rate (%)'. The NBS defines this data as follows "Value-added of Industry refers to the final results of industrial production of t industrial enterprises in money terms during the reference period. Industrial added value can be calculated by two approaches: the

production approach, i.e. gross industrial output value minus intermediate input plus value-added tax, and the income approach, i.e. income for various factors used in the course of production, including depreciation of fixed assets, remuneration of labourers, net of production tax, and operating surplus" (National Bureau of Statistics China, 2020). Figure 3.8 is a plot of the percentage change data from the NBS.

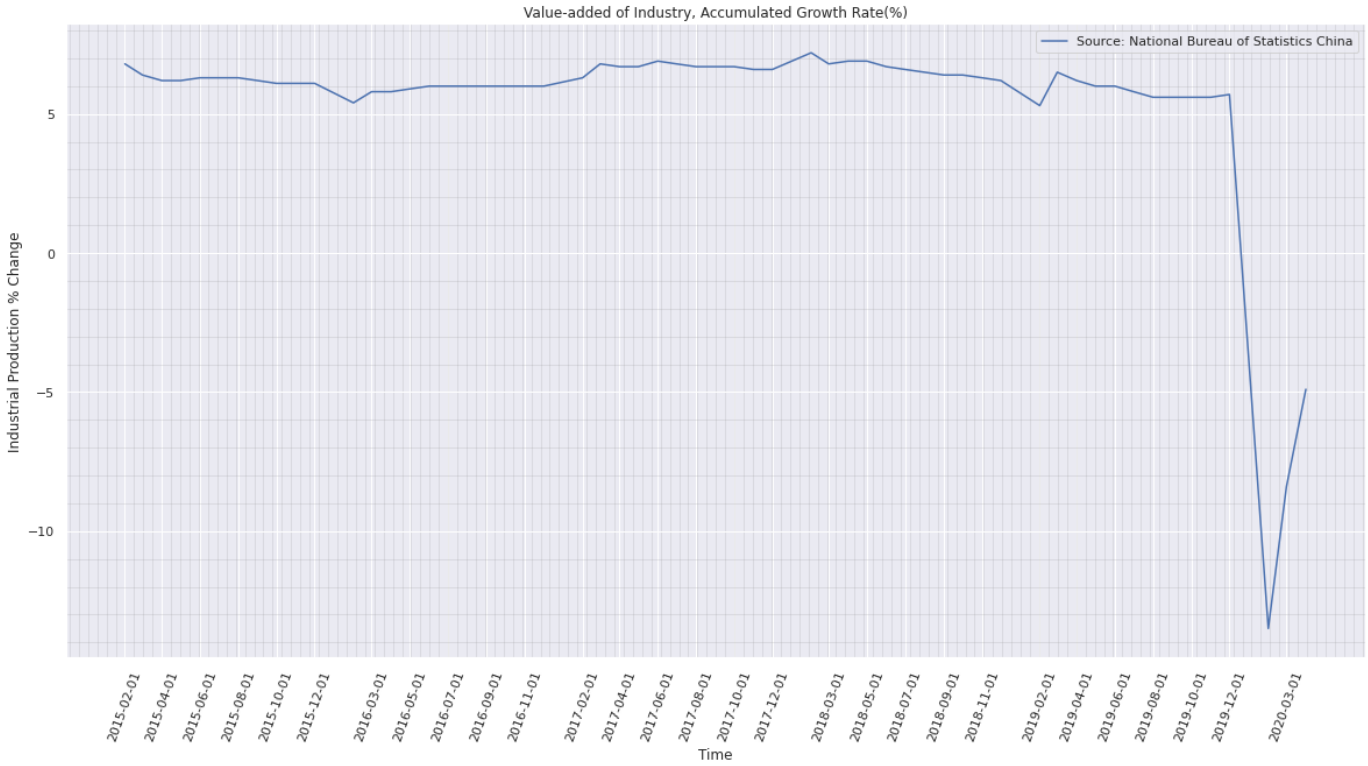


Figure 3.8: Value-added of Industry, Accumulated Growth Rate (%)

This data is only published for the months February-December since according to the NBS "In order to eliminate the impact of the different date of Spring Festival of each year, and enhance the comparability of data, in accordance with the national statistical system, the data in January and February was investigated and released together." National Bureau of Statistics China (2020). In order to ensure compatibility and the same number of observations with the rest of the data, January of each year has therefore been removed from the Luminosity data-set.

3.4 Baseline OLS Comparison to Past Literature

In order to compute a baseline comparison with the results from Henderson et al. (2012), the Luminosity data will be transformed into percentage changes, since the authors estimate an elasticity (percentage change in Luminosity to the percentage change in GDP). Henderson et al. compute a simple average for each period over each country and therefore it is necessary to compute a simple average for China. The first step in this process is to average each months average block (172,800,000 pixels for the full data-set and 1,728,000 for the reduced data) into one single value for the month. The second stage is to calculate the percentage change between these values, since this involves calculating the percentage change, the first month (January 2015) has been omitted. However, this is not an issue since due to the way the NBS data is published from February to December it would have been removed anyway. In Henderson et al. (2012), the independent variable used in the baseline was $ln(lights/area)$. Since their analysis uses panel data (188 countries from 1992 to 2008), it is necessary to control for the differences in the area of the countries. This thesis only looks at China and therefore it is not necessary to control for this variation in land size between countries. Not dividing by land size does not change the estimated coefficient and therefore will not affect baseline comparison to Henderson et al. (2012).

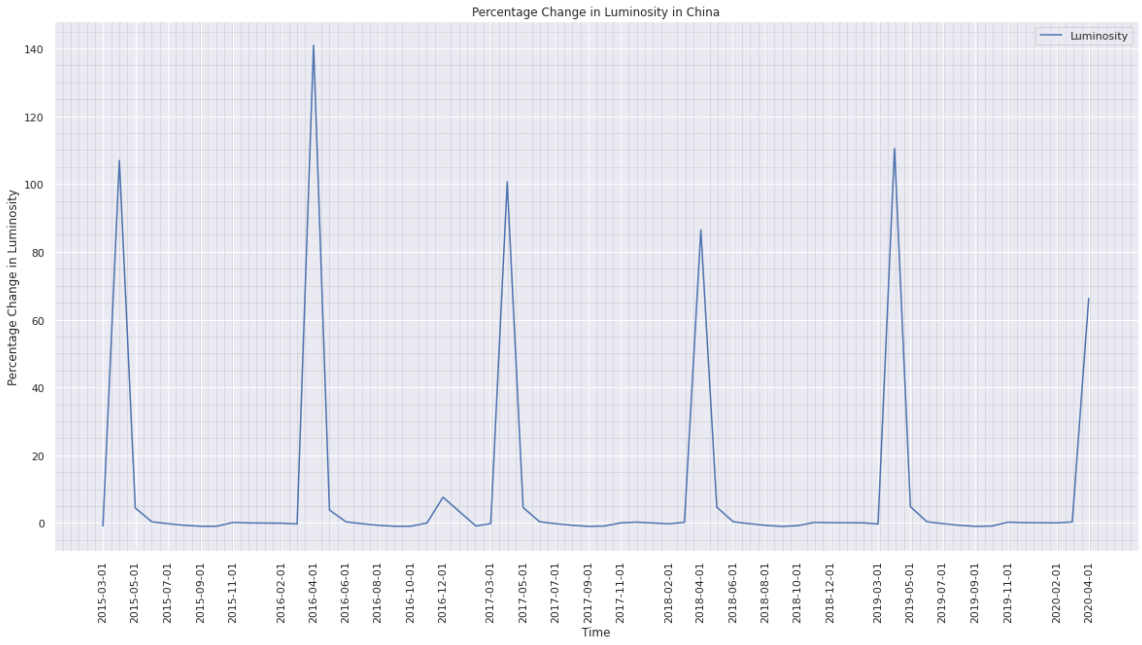


Figure 3.9: % Change in Luminosity Between Months (Jan 2015-April 2020)

Figure 3.9 plots the percentage changes in Luminosity for the period Jan 2015 to April 2020. In April of every year, there is a very large percentage increase in Luminosity of between 80% and 140% (excluding 2020), before returning to being less than 5% for the remainder of the year (excluding December 2016). One potential reason for this could be related to the increased noise in the data when there is a lot of cloud present(shown clearly in Appendix A figure A.2).

However, further research into the data for these months revealed that the issue is not with cloud cover, but instead with a fill value being used. As per Román et al. (2019), there are a range of reasons why a fill value could be used however, the most common are bad quality data or solar zenith angle of less than 108 degrees [5]. Discovering and overcoming this issue is documented fully in appendix C.

3.5 Neural Network

3.5.1 Overview

Although ANN's have only recently gained their huge popularity, they date back to the paper McCulloch and Pitts (1943), who created a stylised methodology to look at the way brains of animals perform computational tasks. Although this framework has been around for a long time, there have been a number of waves of interest in the methodology. However, it is more likely than ever that it will stay in the spotlight in the future for the following key reason:

- Dragland (2013) find that as of 2013, 90% of data had been collected in the period two years prior to the article. This large increase in data collection, commonly referred to as 'Datafication', allows for more in depth analysis and better inference for research tasks.

The structure of the network proposed by McCulloch and Pitts (1943) was by today's standards basic, due to the infancy of the topic. It contained a range of so called 'Biological Neurons' (more commonly referred to as artificial neurons) which contain

[5]The solar zenith angle is the angle between the zenith (the imaginary point directly above the location the image is being taken) and the centre of the sun's disk (the area of light the sun is emanating on the earth).

a number of on/off switches leading to one binary output.

The Neural Network used in this thesis was coded using the Keras library on top of Tensorflow in Python.

3.5.2 Normalisation

Industrial Production has been normalised [6]. Normalisation is the process of re-scaling data to be in a different range. In machine learning (ML) package Scikit-learn, the prepossessing item MinMaxScaler() scales each value to be between 0 and 1 using the following formula from Pedregosa et al. (2011).

$$X_{std} = (X - minX)/(maxX - minX)$$

$$X_{scaled} = X_{std} * (maxX - minX) + minX$$

For NN's, normalization is an important step, since the NN assigns weights to the parameters in the model in order to try and maximise the accuracy (or minimise the loss) of the predictions. Without normalising data, these weights would have to be very large [7]. A model with large weights can be very unstable and have poor predictive power (Brownlee, 2019). Therefore, using the MinMaxScaler() function to normalise the data between 0 and 1 will improve model accuracy and thus the quality and accuracy of the predictions made.

3.5.3 Model Structure

A Deep Neural Network (DNN) is an ANN which contains more than one layer between the input and output layers.
Although most problems can be solved using a single layer, they are less efficient for the following reasons summarised well in (Gron, 2017):

- *Parameter Efficiency* - DNN's can model more complex functions using fewer

[6]Normalisation was included into the percentage change process of the Luminosity data however, it would assign some values to be infinite. This led to the model failing in the fitting process and therefore the percentage change in Luminosity has not been normalised in this thesis.

[7]Especially in the case of this data which has a potentially large outcome variable and training data which can be in a large range (0-65,534).

neurons and therefore they are faster to train.

- Deep networks generalise better to new data-sets.

The type of DNN used in this thesis will be a Convolutional Neural Network (CNN), since images are being used as the independant variable in the model.

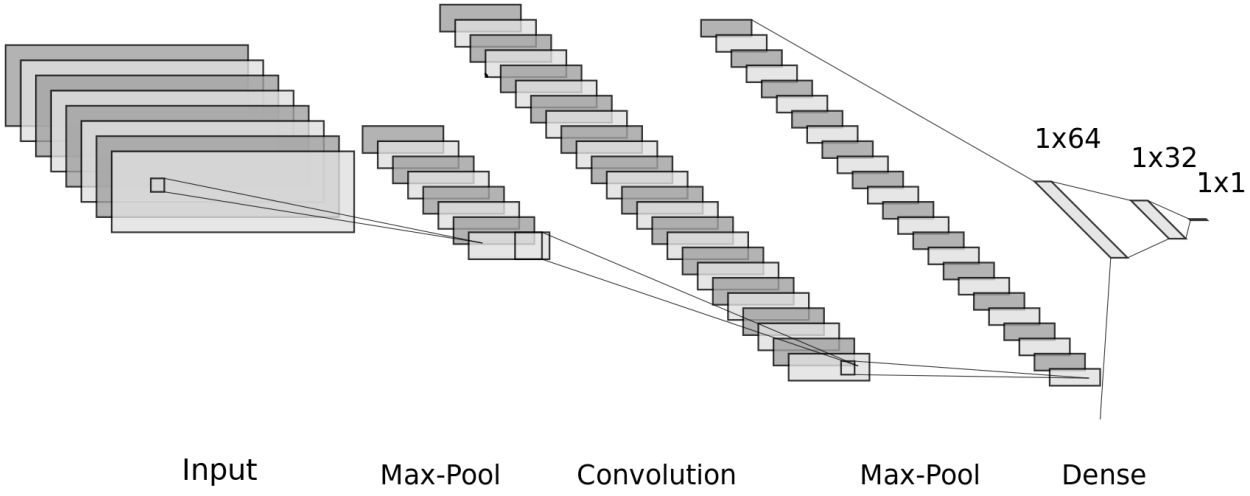


Figure 3.10: Stylised Example of Convolutional Neural Network

Figure 3.10 is a stylised representation of a generic CNN which is fully explained in the proceeding sub-sections. However, as a brief introduction, the input layer in this thesis contains satellite images, the number of these layers can be thought of as the nodes. More layers enables the network to learn more complicated shapes. The Max-Pool layers reduce the dimensions of the model and the Dense layers reduce this down further until the single output shown in the final layer. In this thesis, this output contains the estimate of the percentage change in Industrial Production.

Convolutional layers

At the center of all CNN's are the Convolutional layers, which apply a filter over the input (in this case the satellite images) which gets activated. This filter moves over

the image and outputs a map of these activations called a feature map. Each element of the filter is a weight and when the map is passed over the input the layer calculates the dot product (matrix multiplication of the weights of the filter and the layer) to return the new value. The feature map can detect locations and shapes in the images. In theory, using Convolutional layers should enable the DNN to learn structures of the images. These types of structures include cities and within cities the industrial centres. Using more neurons (which in the context of a CNN neurons are filter maps) enables the network to detect more complicated shapes. For example having one neuron in a layer can allow the network to detect horizontal lines, adding another layer can allow it to detect vertical lines and so on.

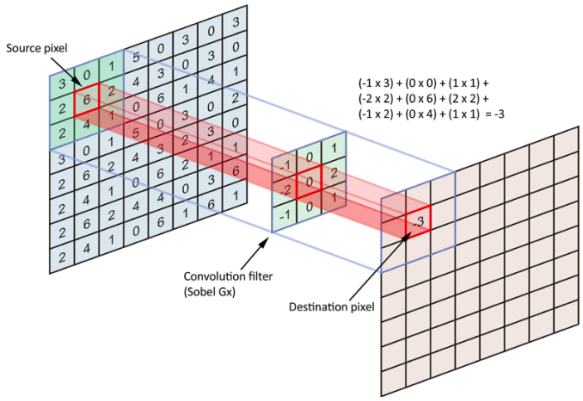


Figure 3.11: Convolutional Layer (Dertat, 2017)

Max Pooling Layers

Max Pooling layers work similarly to Convolutional layers in that they apply a filter, but in this case, the filter is applied to the output of the Convolutional layers. The aim of a Pooling layer is to reduce the dimensions of the layer while keeping the features. Max Pooling works by applying a filter of, for example KxJ. For each KxJ sub matrix it takes the maximum value of this and returns it in the output of the MaxPooling layer that reduces the dimensions of the model. This enables the network to reduce the size of the output of each layer while keeping the detail of the layer.

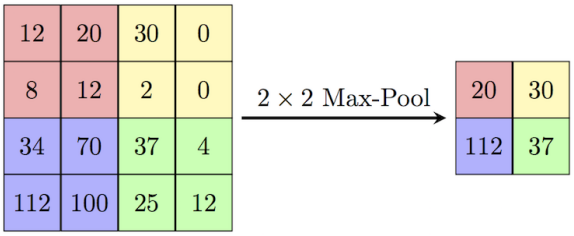


Figure 3.12: Max Pooling Layer (Ricco, 2017)

Hyperparameter Optimisation

There are a number of variables which can be optimised in a DNN.

- *Number of layers* - The number of layers between the input layer and the output layer.
- *Number of neurons* - The number of neurons in each of the hidden layers (can be thought of as the number of filters).
- *Activation Functions* - An activation function determines the output of each of the neurons. They enable the network to have non-linear characteristics.
- *Types of layers* - Convolutional, Max-Pooling, Dense.
- *Optimiser* - Adam optimiser will be used since it combines the benefits Momentum optimisation and RMSProp.

Momentum works by taking a moving average of the gradients [8]. Looking at the moving average of the gradient allows the optimiser to see if the gradient is decreasing (reaching a local or global minima), or increasing (getting further away from the minimum loss). This helps the optimisation to converge to the optimal solution in a faster time.

RMSprop seeks to resolve the problem of trying to minimise losses in a model where the loss function has widely differing gradients. If the learning rate (size of the updates to weights that the NN makes) is too large then the weights go back and forth either side of the optimal weights (at the global minima). It works by dividing the learning

[8]The loss functions of neural networks can be thought of as a 3D map, with the minimum point located somewhere on this map. The steeper the gradient, the further away the model is from the local mimima.

rate by an exponentially decaying average of the square of the gradients. This helps to achieve an optimiser which moves quickly in directions that have a small but consistent gradient and slowly in areas with a large but inconsistent gradient (Hinton et al., 2012).

The number of neurons in the input and output layer are predefined. The input layer is the shape of the array which is fed into the CNN - in this case it is 240x7200 pixels which totals 1,728,000. Since this is classed as a regression problem [9] there is only a single output neuron x where $x \in R$. x can be both positive and negative since, inline with Henderson et al. (2012), the outcome is growth (in this thesis Industrial Production instead of GDP) which, as shown in figure 3.9, takes both positive and negative values.

Within Convolutional layers of the DNN, ReLU activation functions are typically used and can be defined as:

$$ReLU(z) = \begin{cases} 0 & \text{if } z < 0; \\ z & \text{, otherwise.} \end{cases}$$

Gron (2017) find that "in practice it works very well and has the advantage of being fast to compute. Most importantly, the fact that it does not have a maximum output value also helps reduce some issues during Gradient Descent". However, ReLU activation functions are not without fault, since in training some of the neurons can die meaning they always have an attributed weight of 0, if they have a negative number passed to them. However, Xu et al. (2015) find that using one of the so called 'leaky ReLU' activation functions always leads to better accuracy than just using ReLU. Leaky ReLU can be expressed as:

$$LeakyReLU(z) = \begin{cases} a_i z_i & \text{if } z_i < 0; \\ z_i & \text{, otherwise.} \end{cases}$$

where the parameter a_i applies a small positive weight to the negative values and is typically set to 0.01. The authors also found that parametric leaky ReLU (PReLU) significantly outperformed ReLU on large image data-sets (as is the case in this thesis).

[9] Broadly speaking there are two kinds of machine learning problems. Categorical, meaning there are a limited number of outcomes e.g. a number between 1 and 10 in the case of the popular MNIST data-set. The other option is Regression, where the network outputs a single value on a continuous scale.

PReLU differs from Leaky ReLU since instead of the slope coefficient a being predetermined, it allows the NN to optimise this value.

The latest iteration of ReLU is Exponential Linear Unit (ELU) which builds on the benefits of PReLU and Leaky ReLU. ELU activations are closer to zero and tend to be more robust to 'noise'.

$$Exponential\ ReLU(z) = \begin{cases} z & \text{if } z \geq 0; \\ \alpha(exp(x) - 1) & \text{, otherwise.} \end{cases}$$

Where $\alpha > 0$ for ELU.

DNN's, although very effective for prediction, suffer from a range of problems. These include over-fitting and Vanishing/Exploding gradients.

The issue of Vanishing gradients occurs due to back propagation, where weights of the neurons are calculated from the end of the network to the front. Gradients often get smaller as the model progresses to the lower/earlier layers and this can leave the weights attached to these neurons unchanged. This can lead to Vanishing gradients and the training can never converge to a good solution. On the other hand, exploding gradients occur when there are large errors in the training process. Consequently the model makes very large update to the weights of the filters.

Batch Normalisation

Batch Normalisation is another layer which is incorporated into the CNN. In the groundbreaking paper Ioffe and Szegedy (2015), they introduce the Batch Normalisation layer which reduces the Vanishing and Exploding gradients issue referred to above. This methodology works by adding a layer before the activation function. This layer adjusts the weights to be around zero. The layer "lets the model learn the optimal scale and mean of the inputs for each layer." (Gron, 2017). Formally the Batch Normalisation layer works as follows:

$$1. \quad \mu_B = \frac{1}{m_B} \sum_{i=1}^{m_B} X^{(i)}$$

$$2. \quad \sigma_b^2 = \frac{1}{m_B} \sum_{i=1}^{m_B} (X^{(i)} - \mu_B)^2$$

$$3. \quad \hat{X}^{(i)} = \frac{X^{(i)} - \mu_B}{\sqrt{\sigma_b^2 + \epsilon}}$$

$$4. \quad Z^{(i)} = \gamma \hat{X}^{(i)} + \beta$$

- m_B is the number of instances in the mini-batch.
- μ_B is the numerical mean across the mini-batch B.
- σ_B is the empirical standard deviation which is also evaluated across the mini-batch B.
- $\hat{X}^{(i)}$ is the zero-centred and normalised input.
- γ is the scaling parameter for the layer.
- ϵ is a very small number (typically 10^{-3}) in order to make sure there are no issues with division by 0
- $Z^{(i)}$ is the output of the Batch Normalisation operation and is the scaled and shifted version of the inputs.

Ioffe and Szegedy (2015) conclude that including a Batch Normalisation layer improves accuracy in all of the DNN's that they tested and that the vanishing gradients problem reduced significantly.

In machine learning, the data is split into testing and training data-sets. In this thesis, as is typical in many ML problems, a split of 0.8 has been chosen. This means that 80% of the full data-set is used for training and 20% is used for testing. The training data-set is used to fit the model and in this case adjust the weights of the convolutional filters, max pooling filters and dense layers in order to minimise the loss. However, it is important to have a portion of this data-set which is not seen by the model during the fitting process (the training data). This acts as a control to make sure that

the model is not over-fitting. Over-fitting occurs when the model too closely fits a small sample of the overall data, and thus does not generalise well to new data-sets. Therefore, after each epoch (pass through the training data-set) the model is tested with the testing data-set to see how accurate the predictions are. Due to the small number of observations, cross validation (excluding a different subset of the data on each epoch) will not be used. However, if the methodology was extended to include all available years of the Luminosity data (2012-2020), potentially, in this case it would be appropriate to use cross validation.

Chapter 4

Results

4.1 Results

4.1.1 Baseline OLS Results

Table 4.1 contains results for model specifications of regressions which have been run using the reduced data-sets for the explanatory variable (Luminosity). Regression (1) uses the reduced 90% data-set, where the value reported from each 10x10 block was the mean of the block (as per figure 3.4). Regression (2) is also the reduced by 90% data-set, but the value reported from each block is the max value (as per figure 3.5). Regression (3) uses the reduced data-set, but all values over 2000 are set to the mean of the remaining array (as per appendix C). Regression (4) uses data very similar to regression (3), but instead of setting the values greater than 2000 to the mean of the remaining array, it sets these values to 0. The methodology used for the data in regressions (3) and (4) is necessary to try and remove the areas in the summer where data has not been recorded, due to the orientation of the Earth during the summer months in the bright half of the lunar cycle.

The baseline estimate reported in Henderson et al. (2012) is 0.277, which is significant at a 1% significance level (sl). This is an estimate of the elasticity of the change in Luminosity to the change in GDP. This estimate can be intuitively thought of as a 1% increase in the Luminosity results in a 0.277% increase in GDP. The other results in table 4.1 have been calculated using Value-added of Industry, Accumulated Growth

Table 4.1: Dimensionally Reduced Data-set - Baseline OLS Results

	Dependant Variable:				
	% Change in GDP	*Value-added of Industry, Accumulated Growth Rate (%)*			
	Baseline from Henderson et al. 2012	(1)	(2)	(3)	(4)
% Change in Luminosity	0.277***	0.047*	0.049*	3.915	3.537
	(0.031)	(0.025)	(0.025)	(5.274)	(3.503)
Observations	3,015	58	58	58	58
R2	0.769	0.061	0.066	0.01	0.018
Adjusted R2	-	0.044	0.049	-0.008	0.0
Residual Std. Error	-	6.337 (df = 56)	6.32 (df = 56)	6.508 (df = 56)	6.481 (df = 56)
F Statistic	-	3.644* (df = 1; 56)	3.953* (df = 1; 56)	0.551(df = 1; 56)	1.02 (df = 1; 56)
Note:		*p<0.1; **p<0.05; ***p<0.01			

Rate (%) (or in other words Industrial Production) for the outcome variable. It is interesting to compare these estimates since a portion of GDP is made up by Industrial Production. However, although these comparisons are interesting and insightful, it is important to note that this can be likened to an apples to oranges comparison, since they are measuring different metrics.

Regression (1) and (2) are both significant at a 10% sl with estimates of 0.047 and 0.049 respectively. This is approximately 5.8 times smaller than that estimated by Henderson et al. (2012). Regressions (3) and (4) have very large estimated coefficients of 3.915 and 3.537 respectively. Both of these estimates are not significant at the 10% sl. Since these estimates are larger than 1, they are very elastic, meaning that a small change in Luminosity results in a relatively large increase in Industrial Production. The results of regression (3) and (4) vs (1) and (2) suggest that removing the areas of the Luminosity data which have not been recorded in summer is not effective, since the standard errors increased significantly from 0.025 to between 3.5 and 5.3. Therefore, regressions (3) and (4) explain less of the variation of the model and this data will therefore not be used in the following Neural Network section.

Table 4.2: Full Data-set - Baseline OLS Results

	% Dependent variable: Value-added of Industry, Accumulated Growth Rate (%)
	(5)
% Change in Luminosity	0.05**
	(0.025)
Observations	59.0
R2	0.066
Adjusted R2	0.05
Residual Std. Error	6.324(df = 57.0)
F Statistic	4.046** (df = 1.0; 57.0)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4.2 contains the result of the same OLS regression but run with the full data-set. Due to OLS being computationally less expensive than NN's this was possible. The same methodology was applied where for each month's Luminosity data, the average of

all pixels was used as the explanatory variable. This regression estimates an elasticity of 0.05 which is significant at a 5% sl. This is the only estimate presented in this thesis which is significant at the 5% sl. 0.05 is larger than the estimates from regressions (1) and (2) and is much smaller than the estimates from (3) and (4), which seek to correct the blank pixels. In addition, it is 5.54 times smaller than the estimates from Henderson et al. (2012).

Using the point estimate of the elasticity of the change in Luminosity to the change in output from Henderson et al. (2012), it is possible to estimate the percentage change in GDP in China during the first quarter of 2020. In order to calculate the percentage change in Luminosity, the Luminosity by pixel for each month of Q4 of 2019 was averaged to make a quarterly Luminosity array. The mean of this array was then calculated. The same was completed for Q1 of 2020. The percentage change between these two averages (using the full data set) is 82.4121 % which is a surprisingly very large increase. This result deserves further research in the future, as it was surprising to find such a large percentage increase in Luminosity between Q4 2019 and Q1 2020. This could potentially imply that the accuracy of the econometric approach taken by Henderson et al. (2012), when using this data-set for China in this time period, is not applicable.

Using the proposed equation for predictive purposes from Henderson et al. $z_i = \hat{\psi} x_j + e_j$, the estimated percentage change in GDP in China in the first quarter of 2020 is

$$z_{China} = 0.277 \cdot 82.4212\%$$

(where z_{China} is the estimated percentage change in GDP). This yields an estimated increase in GDP in China of 22.83 %. This estimate is higher than was expected due to the percentage decrease in Industrial Production presented by National Bureau of Statistics China (2020), which shows a sharp decrease in Industrial Production. Although, changes in Industrial Production do not necessarily lead to changes in GDP, a large decrease in Industrial Production is likely to be mirrored by a decrease in GDP.

4.1.2 Neural Network Results

A range of different NN specifications were tested, with different model depths (the number of layers in the DNN), number of nodes in each layer, kernel size, Max-Pool layer size, dropout and activation functions. Two specifications with the lowest loss are reported in this section of the thesis, with the full specifications in appendix D. One model specification combining two different input images has also been tested. This model utilised the power of the Keras to input both the reduced mean data (as per figure 3.4) and the maximum reduced data (as per figure 3.5) before combining for one output. The Luminosity data used in this section (for both the single image and part of the double image input models) was the reduced-by-mean data as explained in section 4.1 and shown in figure 3.4.

In addition to the model summary (containing details on model architecture), graphs containing the training and validation loss are included. When looking at the graphs, it is important to compare the training and validation loss curves. If the training loss curve continues to decrease, while the validation loss curve stays constant, then this is a sign of over-fitting.

It is common to use half the number of nodes in the first hidden layer than the number of pixels in the input layer. However, due to the large size of the data used, it was not possible to use this many nodes. This would require 864,000 neurons in the first layer and when tested with a model with a moderate number of layers, it was found that the maximum number of neurons that could run on Google Colab was only 256.

Model With Two Image Inputs

Figure D.1 in appendix D is a representation of the specification used in the CNN model with two image inputs. The reason two image inputs were used is that potentially, combining the results of the dimension reducing methodologies similar to 'Max Pooling' and 'Average Pooling', will enable the model to gain more insight into how both the maximum and average values change with changes in Industrial Production. This may lead to increased accuracy in the predictions. A number of different epochs and models have been tested, but the results presented below show the most favourable model (lowest loss with the minimum number of epochs). In these models 20 epochs

were used but even the model with 100 epochs did not reduce the loss significantly.

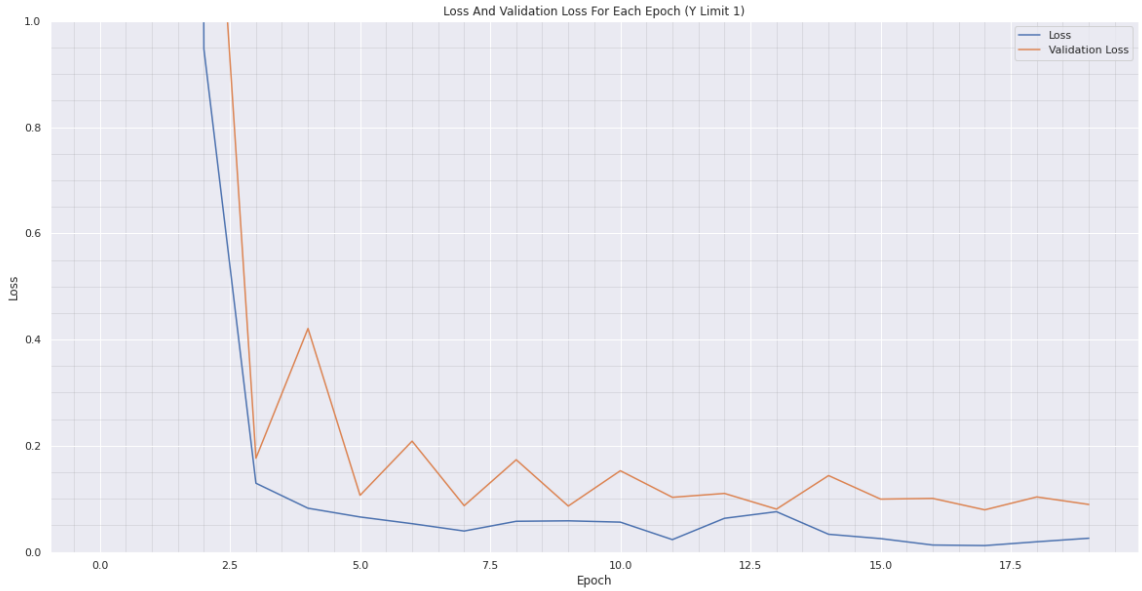


Figure 4.1: Loss and Validation Loss For Model With Two Image Inputs (Y Axis Limited to 1)

Figure D.3 and 4.1 show the loss [1] (from the training process) and validation loss (from the validation step at the end of each epoch). Loss decreases significantly during epoch 1 and then consistently decreases and stabilises throughout the remaining epochs. Validation loss, however, fluctuates between 0.08 and 0.42 (as shown more clearly on figure 4.1). This relatively low validation loss suggests that the model is effective in predicting the percentage change in Industrial Production from a percentage change in Luminosity.

Model With One Image Inputs

Figure D.2 in appendix D is a representation of the specification used in the CNN model with one image input. This specification uses the reduced data-set where the dimension reducing process followed the methodology set out in section 3.2.1 and shown in figure 3.4.

[1] measured by the mean square error

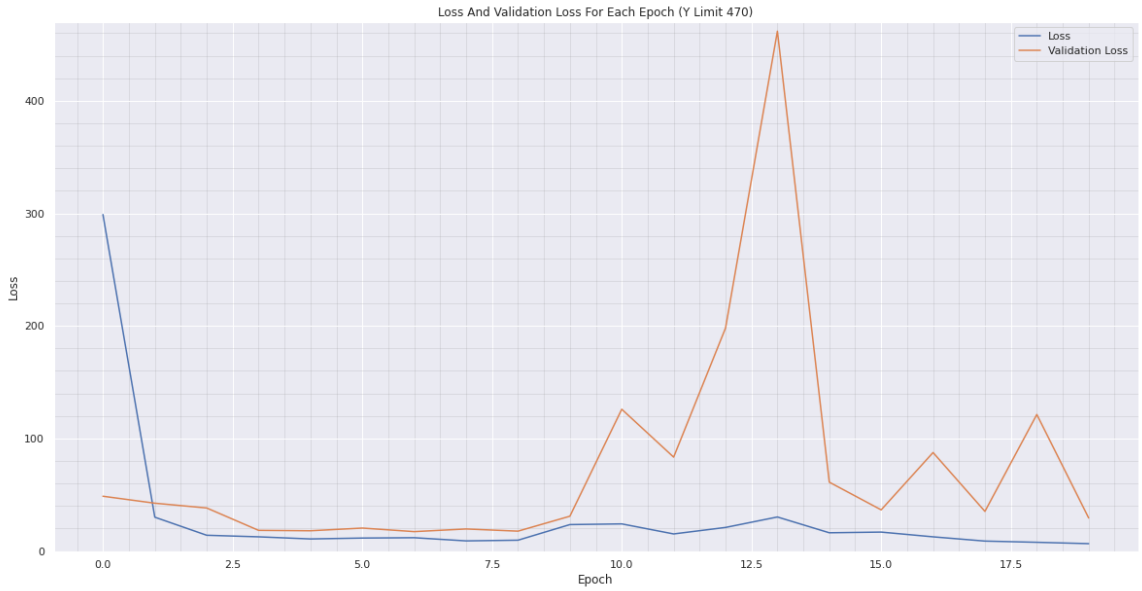


Figure 4.2: Loss and Validation From Single Image Model

Figure 4.2 shows the large and unstable loss and validation loss which the model with only a single image produces. This is significantly inferior in terms of accuracy when compared to the two image input model. The loss stays consistently around 20 until epoch 9 where it increases considerably to 470 before decreasing back to less than 120. This large variance in loss does not suggest that the model is effective in its predictions.

Figures 4.1 and 4.2 (D.3 shows the full y-axis range of values for the two image input model) show that both loss and validation loss are significantly smaller in the model with two input images, suggesting that it has more explanatory power.

Using A CNN Model To Make Predictions

Although the training data-set will need to be reduced in the number of observations (by 3), it is interesting to see the predictions made by the CNN. Since the two image input model had a far lower loss, this is the model which has been been selected for these predictions. The Luminosity data for February 2020 to April 2020 was removed and the model was retrained for the period February 2015 - December 2019 (February was the first observation due to Chinese New Year festival, details of this are covered in section 3.3).

The model was subsequently evaluated using the Luminosity data from February 2020 to April 2020, with the following predictions. A graph of these predictions (figure 4.3) is presented below:

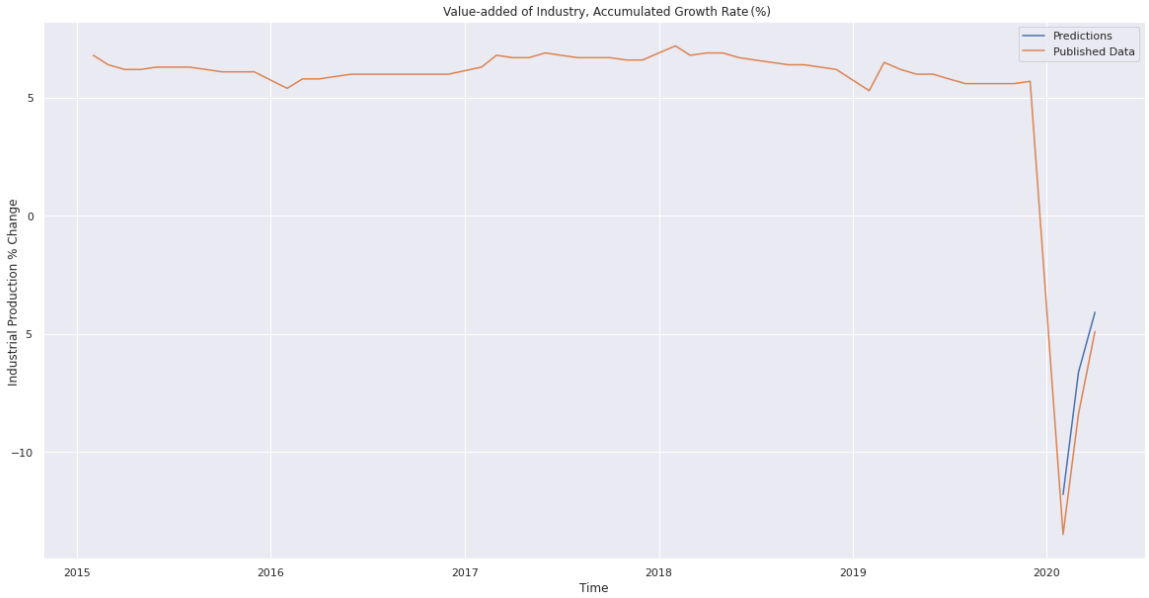


Figure 4.3: CNN Predictions With Two Image Model

Figure 4.3 plots the published percentage change in Industrial Production in China (National Bureau of Statistics China, 2020) as well as the predictions from the Convolutional Neural Network. These predictions from the CNN accurately predict the change in Industrial Production in the case of China for this period. This is a very promising result, since the data used was reduced in size by 90%. Given the full data-set, it is likely that these predictions would be even more accurate.

Chapter 5

Discussion and Conclusion

5.1 Discussion and Conclusion

Despite the promising results, this thesis alone cannot not provide a definitive answer as to whether Neural Networks are effective in predicting Industrial Production. In order to get closer to answering this question, further research would need to be completed. Firstly, replicating this analysis for a range of countries would lead to a better and more complete picture as to whether this methodology is effective. Since this thesis only evaluates the model in the case of China, the low loss and close predictions may be spurious. A range of different countries with different characteristics should be used. Some examples of how these characteristics can vary include, but are not limited to;

- Population
- Land size
- Employment sector composition
- GDP per capita
- Cultural and religious groups
- Urban to rural living ratio (proportion of people living in urban areas to rural areas)

Using the elasticity formalised in Henderson et al. (2012) to estimate the decrease in GDP in China (results in 4.1.1) in Q1 2020, led to some doubt in the accuracy of this elasticity when used with the NPP data. Since the National Bureau of Statistics of China reported a decrease in Industrial Production for all months in Q1 2020, it is unlikely that this would have been complemented with an increase in GDP of 22.83% over the same period.

Repeating the analysis with the DMSP-OLS data, used in both Henderson et al. (2012) and Bickenbach et al. (2016), instead of the Suomi NPP data would lead to an insightful comparison as to which provides the better predictions in the case of NN's. Although Dai et al. (2017) found that the NPP data was more effective when finding an elasticity using econometrics, it may be that the less detailed data with reduced noise leads to more accurate predictions from the Neural Network with lower loss.

The CNN's used in this thesis have been trained on a significantly reduced data-set. Therefore, it is likely that if the full data-set had been used, the Neural Network would be able to extract more information about the change in Luminosity and therefore would be able to yield more accurate predictions. In order for this to be possible, having access to a Server or Workstation computer with significantly more RAM and Nvidia Graphics Cards (GPU) [1] would increase processing power substantially. Even the reduced data-set (approximately 2GB in size), when fed into the DNN, with a relatively small number of layers and nodes, uses 32GB of RAM. Therefore, if there was access to a server with approximately 2TB or more of RAM this should enable the modelling of the full data-set. The full data-set has 100 times more pixels and consequently it is likely that the data will contain more nuance and information than the reduced data-set and thus should produce increased accuracy in the predictions.

Using a substantially more powerful computer would lead to the further benefit of being able to use a more complex model. More computational resources would allow the model to have more layers (and thus extract more detail). It would also allow more nodes per layer which would result in the model being able to recognise more complicated shapes, for example industrial areas and resource processing areas. Therefore, us-

[1] Nvidia GPU's in particular can be used to share the processes of Deep Learning with the CPU, as well as having faster RAM.

ing a more powerful computer and subsequently a more complicated model (especially when more data-sets are included in the model) could yield very large improvements in the model.

With access to more computing power, the full data-set could be run (the full data-set being both complete in terms of dimensions as well as having data from January 2012 to the present). This increase in both the information per observation and the number of observations may allow the NN to be trained with GDP as the outcome variable. This would allow direct comparisons to be made with the framework introduced by Henderson et al. (2012) which in turn would enable direct comparisons of the accuracy and loss of the frameworks. Subsequently, it would also be interesting to compare which framework is more effective at predicting changes in GDP in more extreme circumstances, as has been the case in the COVID-19 crisis. Once these comparisons have been made (along with the addition of analysis for a range of countries), it will then be possible to draw more concrete conclusions on the effectiveness of Luminosity data and CNN's in predicting changes in Industrial Production (and GDP).

Since the inclusion of a second image input significantly reduced both loss and validation loss, including other metrics, for example pollution, could be a further extension of the model. This could be used to improve accuracy of the predictions. As per the findings from Liu (2006) (covered in 2.1), there is a Granger causal relationship between Nitrogen Oxides $ln(NO_x)$ and $ln(Y)$. Thus, including this data in the NN may help to improve the accuracy of the predictions. In the case of China where the economy has a lot of Industrial Production (and therefore pollution), it may add a large amount of insight to the changes in production, especially in the case of the COVID-19 crisis where production decreased significantly as the factories were closed due to lock-down measures. Using NOx data from the Sentinel-5P satellite would be applicable for this extension. It was not included in this thesis since the data-set is not split up by country and would have been very complicated and time consuming to process.

Once the VNP46A2 data-set is released, this data could potentially be used instead. As mentioned in section 3.2, VNP46A2 has the benefit that it has been processed by NASA to remove cloud cover and other environmental noise like Aurora. Figure

3.1 shows the noticeable improvements in clarity bought by VNP46A2. In the case of China (illustrated in figure A.2 in appendix A), there can be a lot of cloud cover, which in turn reduces or blocks the radiance of the Luminosity from the industrial areas. This results in a very noisy signal for the Neural Network and will reduce accuracy. One of the large benefits with VNP46A1 however, is its short delay between the capture date and the publish date (due to not being heavily processed). This enables analysis to be produced shortly after the event. In contrast, for example, VNP46A2 data captured in 2012 and 2013 will be released in Summer 2020.

As discussed in section 3.3, there has been a decrease in Industrial Production to GDP ratio in recent years. Therefore, in the future, if this methodology continues to be used and applied to other countries, it may be preferable to use a proxy more based upon output of the services industries, for countries in a similar stage of development and employment mix to China. In addition, industries in the service sector are likely to be heavily office based and therefore will emanate a lot of light into the atmosphere.

An issue with using Neural Networks instead of the estimated elasticity proposed by Henderson et al. (2012) is transferability. Although the findings in Jean et al. (2016) show that using a model trained in one country, to make predictions about another country shows good predictive power this is not easily possible with the methodology proposed in this thesis. When using a CNN, it is not possible to use a different image size as the input to the model when fitting the model and when making predictions. This means that, for example, it is not possible to use the model trained on China to make predictions for Italy, due to the different land size. In addition, it is likely that the weights of the model would be different. It would be naive to assume that the effects of changes in lights to changes in Industrial Production or GDP are the same between different countries. In the case of an economy such as China with one coast, a large proportion of this economic activity occurs in the provinces on this coast. Consequently, it is to be expected that an increase in economic activity are likely to be seen most in these provinces. Therefore, for the filters of the NN, as it passes through areas inland it will not become activate as these coastal provinces. In comparison, the USA (the closest country in terms of land size) has a South, East and West Coast. When looking at Luminosity images for the USA (as shown in Chen and Nordhaus

(2011)), the economic activity is spread more evenly around the country. With centres close to both coasts (particularly in the States surrounding the East Coast) as well as by the Great Lakes in the North. These different geographical structures would be likely to require different weights to minimise loss associated with the predictions.

This methodology has a wide range of policy applications for both the private and public sectors. Firstly, international institutions such as the World Bank or the International Monetary Fund may use this methodology when completing their analysis on nations which rank poorly on the Penn World Tables (PWT). These are countries for which the official national statistics have the problems of poor accuracy and irregular publishing frequency. For example, if the World Bank was assessing a request for a loan and they were sceptical about the accuracy of the GDP or Industrial Production statistics presented to them by the country. They would be able to complete their analysis independently and verify any statistics before making their decision. If further testing on the methodology presented in this thesis suggested or found it to be more accurate than the existing econometric approach, their analysis would have greater accuracy and a more informed decision could be taken.

Although this brings large benefits, it also comes with costs. For governments in developing countries with very low budgets, they may not be able to afford the computing power for the computationally expensive task of running a Convolutional Neural Network. As such, the traditional econometric approach is much more accessible to these nationals despite the potentially lower accuracy. If this was the case then it could lead to even greater issues in these countries. Without accurate prediction of key statistics such as Industrial Production or GDP, it may prove difficult for the government in question to allocate their funds in the most productive way.

To conclude, this thesis developed a methodology for using satellite images of Luminosity and Convolutional Neural Networks to predict changes in Industrial Production. The results presented in section 4.1.2 and in particular, the two image model (full model presented in the figure D.1 in appendix D) produces low losses and accurate predictions. Figure 4.3, which is trained on a smaller sample to allow predictions for 2020 to be made, matches closely with the true reported percentage change in Industrial Production from the Chinese NBS during the COVID-19 crisis. This thesis has

also discussed a range of potential improvements and research suggestions. These can be carried out to further develop a complete picture as to whether Neural Networks are truly effective in predicting economic indicators such as Industrial Production and GDP and how this compares to traditional econometric approaches. Consequently, the results from this thesis suggest there is considerable scope for further research in this field in the future.

Bibliography

BBC News. (2020). A quick guide to the us-china trade war. BBC. https://www.bbc.co.uk/news/business-45899310

Berliant, M. & Weiss, A. (2017). Measuring economic growth from outer space: A comment.

Bickenbach, F., Bode, E., Nunnenkamp, P. & Söder, M. (2016). Night lights and regional gdp. *Review of World Economics*, *152*(2), 425–447.

Brownlee, J. (2019). How to use data scaling improve deep learning model stability and performance. https://machinelearningmastery.com/how-to-improve-neural-network-stability-and-modeling-performance-with-data-scaling/

Campbell, D. (2020). *Uk coronavirus crisis 'to last until spring 2021 and could see 7.9m hospitalised'*. https://www.theguardian.com/world/2020/mar/15/uk-coronavirus-crisis-to-last-until-spring-2021-and-could-see-79m-hospitalised

Carlowicz, M. (2020). Nighttime images capture change in china. NASA. https://earthobservatory.nasa.gov/images/146481/nighttime-images-capture-change-in-china

Chen, X. & Nordhaus, W. D. (2011). Using luminosity data as a proxy for economic statistics. *Proceedings of the National Academy of Sciences*, *108*(21), 8589–8594.

Dai, Z., Hu, Y. & Zhao, G. (2017). The suitability of different nighttime light data for gdp estimation at different spatial scales and regional levels. *Sustainability*, *9*(2), 305.

Department of Health. (2020). Covid-19: Epidemiology, virology and clinical features. Department of Health. https://www.gov.uk/government/publications/wuhan-novel-coronavirus-background-information/wuhan-novel-coronavirus-

epidemiology-virology-and-clinical-features#:~:text=On%2031%20December%202019,,Hubei%20Province,%20China.

Dertat, A. (2017). Applied deep learning - part 4: Convolutional neural networks. Towards Data Science. https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2

Dragland, Å. (2013). Big data, for better or worse: 90%; of world's data generated over last two years. https://www.sciencedaily.com/releases/2013/05/130522085217.htm

Gron, A. (2017). *Hands-on machine learning with scikit-learn and tensorflow: Concepts, tools, and techniques to build intelligent systems* (1st). O'Reilly Media, Inc.

Henderson, J. V., Storeygard, A. & Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, *102*(2), 994–1028. https://doi.org/10.1257/aer.102.2.994

Hinton, G., Srivastava, N. & Swersky, K. (2012). Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. *Cited on*, *14*(8).

Ioffe, S. & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B. & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science*, *353*(6301), 790–794.

Liu, G. (2006). *A causality analysis on gdp and air emissions in norway* (tech. rep.). Discussion Papers.

McCulloch, W. S. & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, *5*(4), 115–133.

NASA. (2020). The vnp46a2 daily moonlight-adjusted nighttime lights (ntl) product. NASA. https://blackmarble.gsfc.nasa.gov/VNP46A2_RMS.html

National Bureau of Statistics China. (2020). 'value-added of industry, accumulated growth rate(%)' [data retrieved from World Development Indicators, http://data.stats.gov.cn/english].

OECD. (2020). Industrial production (indicator). https://doi.org/10.1787/39121c55-en

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Ricco, J. (2017). What is max pooling in convolutional neural networks? Quora. https://www.quora.com/What-is-max-pooling-in-convolutional-neural-networks

Román, M. (2020). Viirs/npp daily gridded day night band 500m linear lat lon grid night - laads daac. NASA. https://ladsweb.modaps.eosdis.nasa.gov/missions-and-measurements/products/VNP46A1

Román, M. O., Wang, Z., Shrestha, R., Yao, T. & Kalb, V. (2019). Black marble user guide version 1.0. *NASA: Washington, DC, USA*.

World Bank. (2018a). Exports of goods and services (current us$) [data retrieved from World Development Indicators, https://data.worldbank.org/indicator/NE.EXP.GNFS.CD?most_recent_value_desc=true].

World Bank. (2018b). Gdp (current us $) [data retrieved from World Development Indicators, https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?most_recent_value_desc=true].

World Bank. (2018c). Industry (including construction), value added (% of gdp) [data retrieved from Global Economic Monitor, https://data.worldbank.org/indicator/NV.IND.TOTL.ZS].

World Bank. (2020). Global economic monitor [data retrieved from Global Economic Monitor, https://databank.worldbank.org/source/global-economic-monitor-(gem)].

Xu, B., Wang, N., Chen, T. & Li, M. (2015). Empirical evaluation of rectified activations in convolutional network. *arXiv preprint arXiv:1505.00853*.

Appendix A

Appendix A - Satellite Images

h25v03	h26v03	h27v03	h28v03	h29v03	h30v03	h31v03
h25v04	h26v04	h27v04	h28v04	h29v04	h30v04	h31v04
h25v05	h26v05	h27v05	h28v05	h29v05	h30v05	h31v05
h25v06	h26v06	h27v06	h28v06	h29v06	h30v06	h31v06
			h28v07	h29v07		

Table A.1: Stylised Representation of Geographical Location of Each Block Used in the Satellite Images.

Figure A.1: 25^{th} February 2020 Night Lights over China - Low Noise

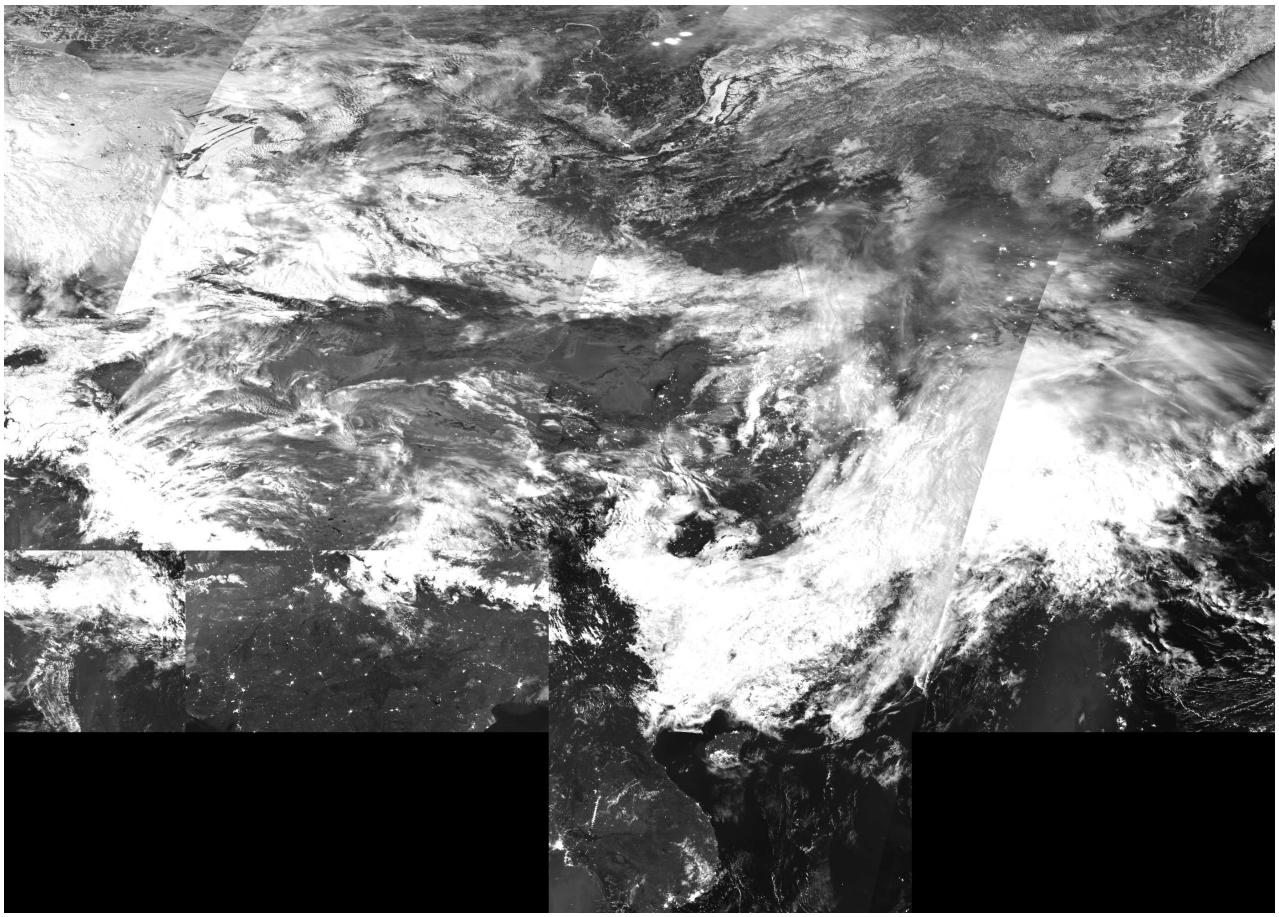


Figure A.2: 19^{th} February 2020 Night Lights over China - High Noise

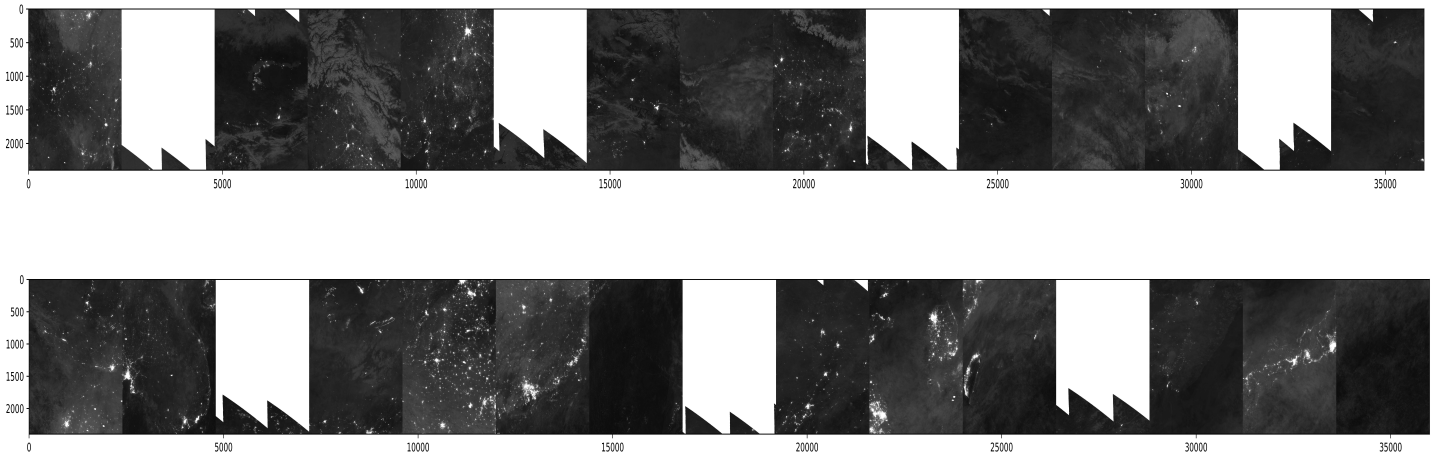
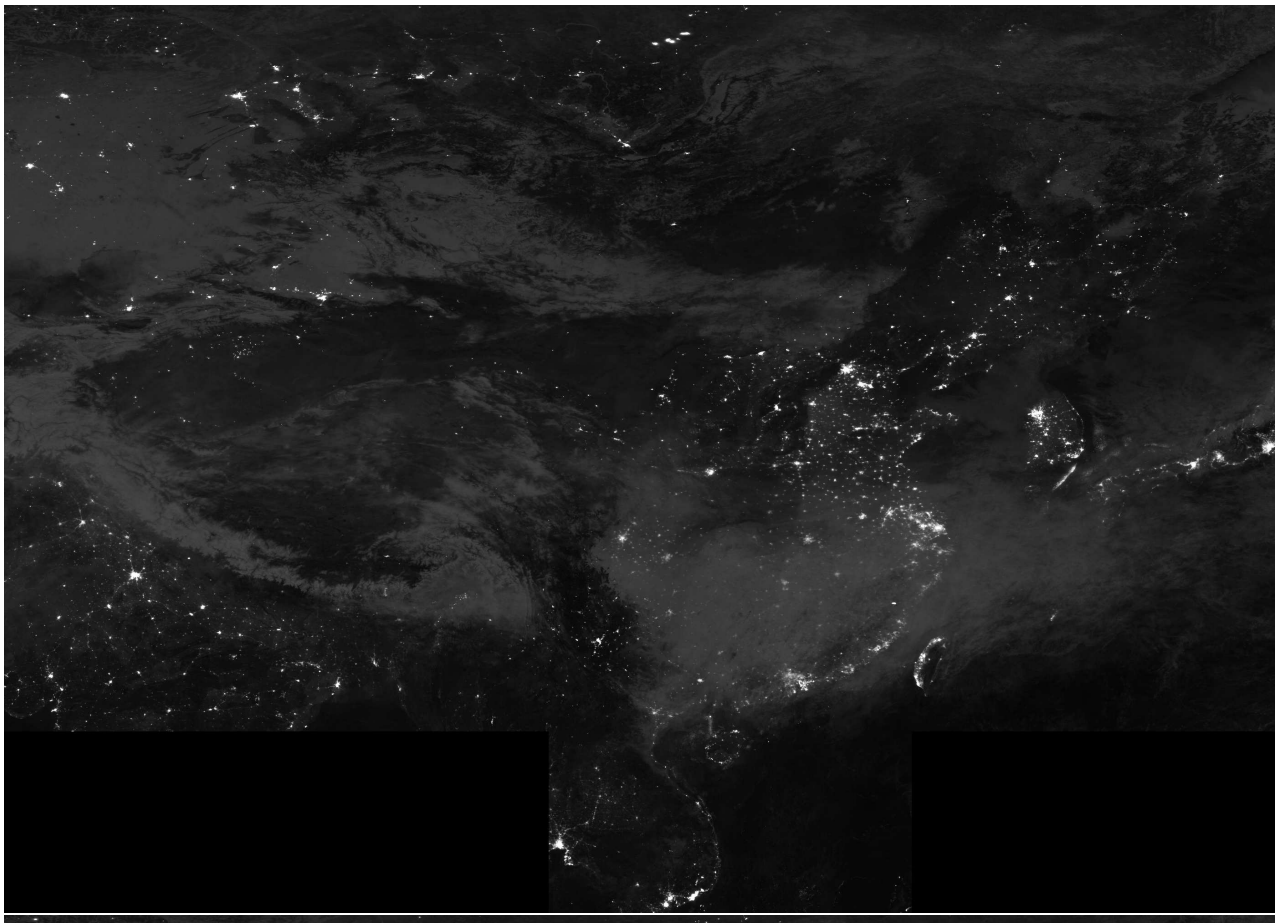


Figure A.3: VNP46A1 Data Concatenated

The data is split over two images to make it easier to display within this document.

Figure A.4: Averaged February 2019 Night Lights over China



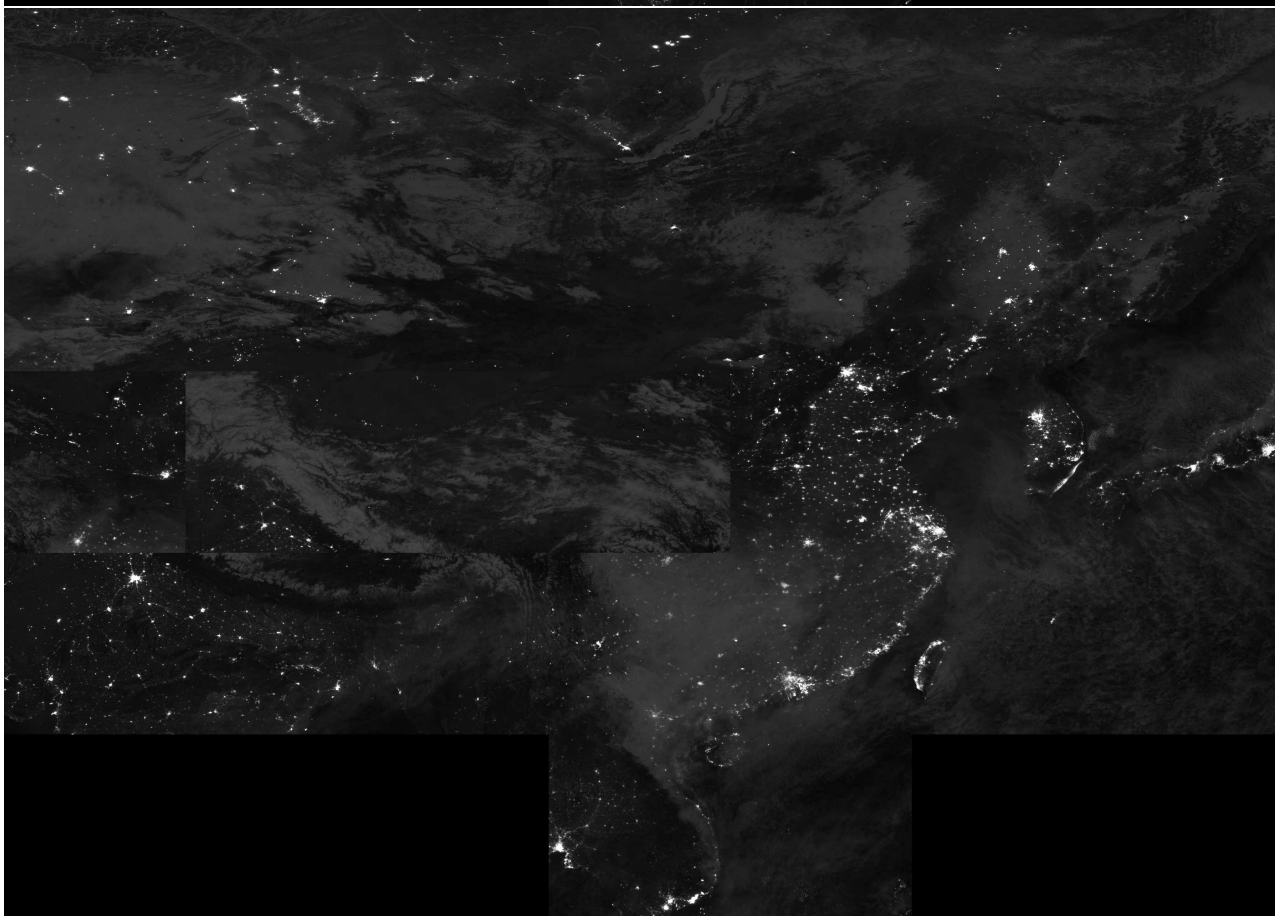


Figure A.5: Averaged February 2020 Night Lights over China

Appendix B

Appendix B - Descriptive Statistics And Graphs

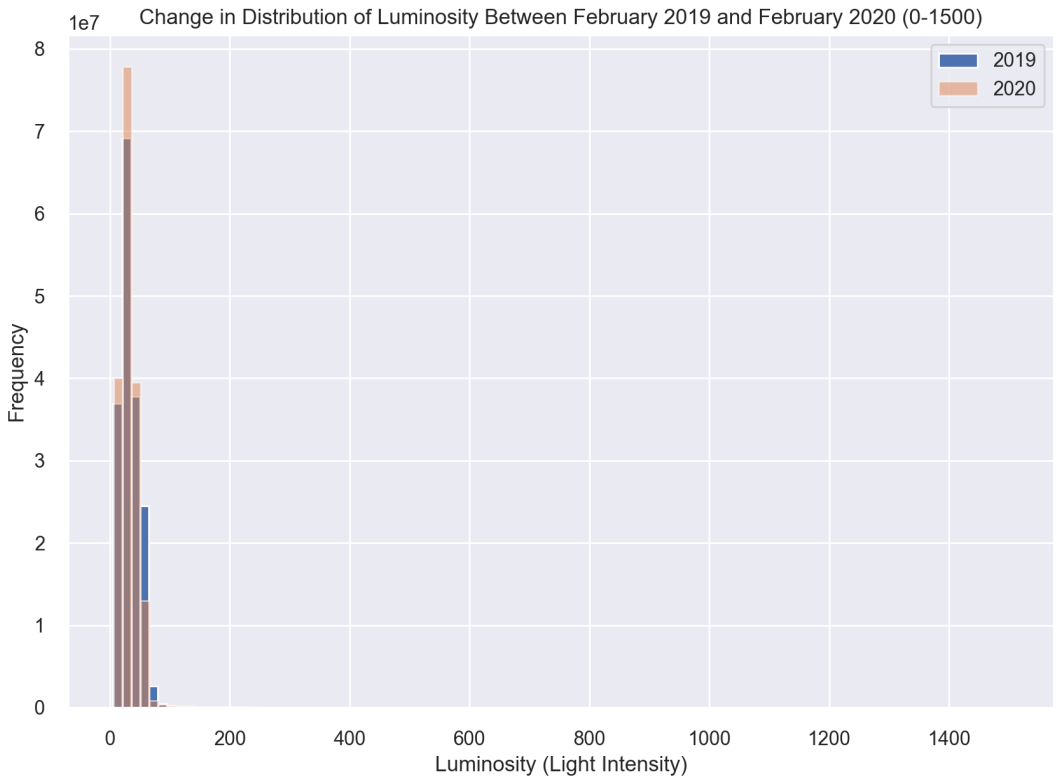


Figure B.1: Histogram Comparing February 2019 and February 2020 Luminosity Data (Range 0-1500).

	10% Feb 2019 Avg	*10% Feb 2019 Max*	*Full Feb 2019*	*10% Feb 2020 Avg*	*10% Feb 2020 Max*	*Full Feb 2020*
Count	1728000	1728000	172800000	1728000	1728000	172800000
Mean	33.90797	45.54164	33.90797	32.96723	45.51349	32.96723
STD	30.60179	140.31260	40.82944	27.14158	131.35510	36.97101
Min	5.79536	6.32143	4.92857	7.38172	8.34483	6.48276
25%	21.51179	26.10714	21.14286	22.24552	26.93103	21.89655
50%	30.06536	36.57143	29.85714	29.06310	36.13793	28.82759
75%	43.15321	50.85714	43.21429	40.07896	48.13793	40.20690
Max	8011.04932	33390.53516	33390.53571	7140.84277	33686.00000	33686

Table B.1: Descriptive Statistics of Full vs Reduced Data-sets

Methodology	*Shape*	*Unique Values*
Original full array.	2400x72000	52,093,631,722
Normalising values in range 0-255 then exporting as PNG file.	2400x72000	108 [1]
Saving full array as PNG then importing intro Keras.	2400x72000	117
Averaging each 10x10 array to return one value.	240x7200	391,112
Returning the maximum value of each 10x10 array.	240x7200	35,527

Table B.2: Comparing The Number Of Unique Values And Shape Between The Full Array And A Number Of Reducing Techniques For April 2015

Table B.2 shows that taking the average of the array and reducing every 10x10 pixel block, provides by far the largest number of unique Luminosity values. It produced over 10 times the number of unique values compared to taking the maximum and approximately 3555 times as many values as when the arrays were exported as images. Therefore, it is very likely that using the average will produce the best predictions since the data contains the largest amount of variation. Both the Max and Average values will be tested in the analysis.

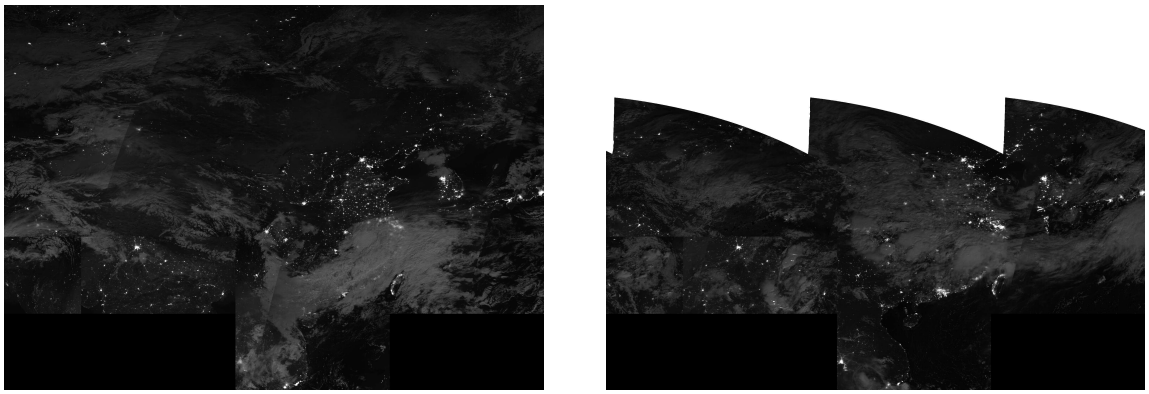
[1] The value in cell of figure B.2 is smaller than the value below (117) because this methodology normalises the data, which reduces the number of unique values, before saving as a PNG.

Appendix C

Appendix C - Removing Blank Areas In Luminosity Data

Section 3.4 uncovered an issue (shown in figure 3.9) where every April, there is a very large percentage increase in Luminosity. Initially, it seemed that this issue may have been the result of entering the rainy season (more clouds lead to higher values in the data since the cloud is very light in colour).

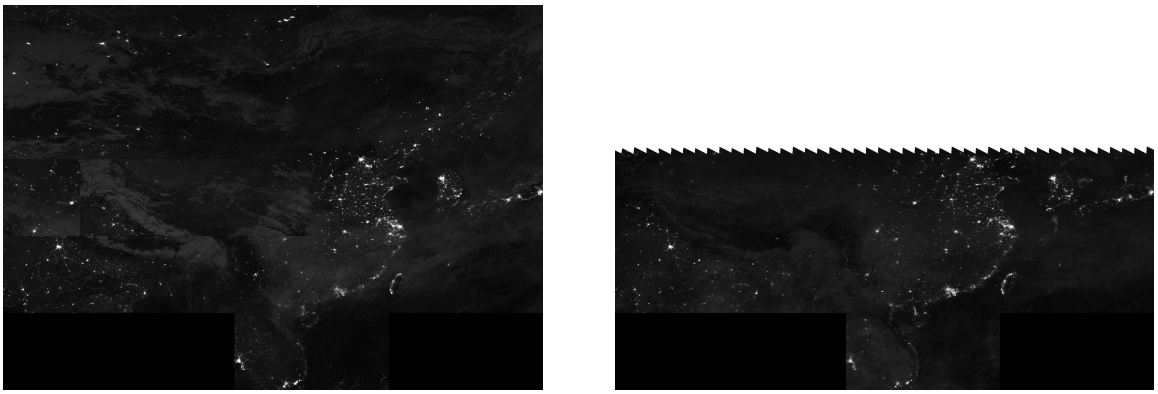
However, further research into the data for these months revealed that the issue is not exclusively due to cloud cover, but instead is significantly influenced by the area that the Satellite records its data, due to being in the bright half of the lunar cycle. This is well illustrated in the images in figure C.1. The 21st of June has been selected since this is the longest day of each year, it is therefore the most extreme example of the cropping of the data.



(a) 24^{th} March 2019 (b) 21^{st} June 2019

Figure C.1: NASA not Recording Data For All of China

This cropping is also present in the averaged data.



(a) March 2019 Average (b) June 2019 Average

Figure C.2: NASA not Recording Data For All of China (Averaged over the Full Month)

These images help to explain the likely reason for this large increase in the average Luminosity (shown as white pixels in the images), since for these areas where the Luminosity data has not been recorded, the fill value 65,535 is used.

In order to overcome this issue, during the dimension reducing process (covered in section 3.2.1), a new array (referred to as array B) was created where all of the values (of the original array A) above 2,000 [1] were set to 0. The mean of array B was then

[1]Originally the cutoff value was set to 65,534 and later 40,000 but this did not remove the blank areas. Potentially, this is due to the averaging process where in each month different days have different areas which are blank. However, setting the value to 2000 was successful in removing these blank spots.

taken ($\bar{B}$) and for all of the values in array A which were greater than 2,000, they were changed to $\bar{B}$. In addition to this methodology, a third version of the averaging methodology was used, which set all values larger than 2000 equal to 0. However, this methodology of setting the values to 0 may reduce the amount of Luminosity indicated by the data compared to months where Luminosity is recorded for the full geographical area. Therefore, for every April the percentage change may be unrepresentatively small and therefore could potentially lead to underestimates of the increase in Luminosity.

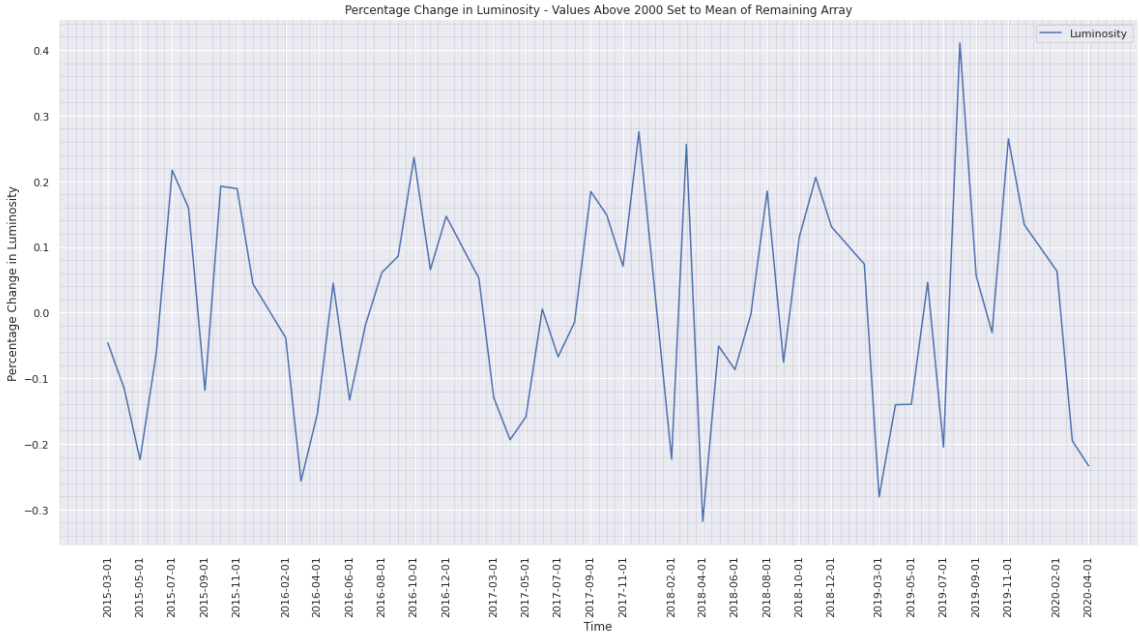
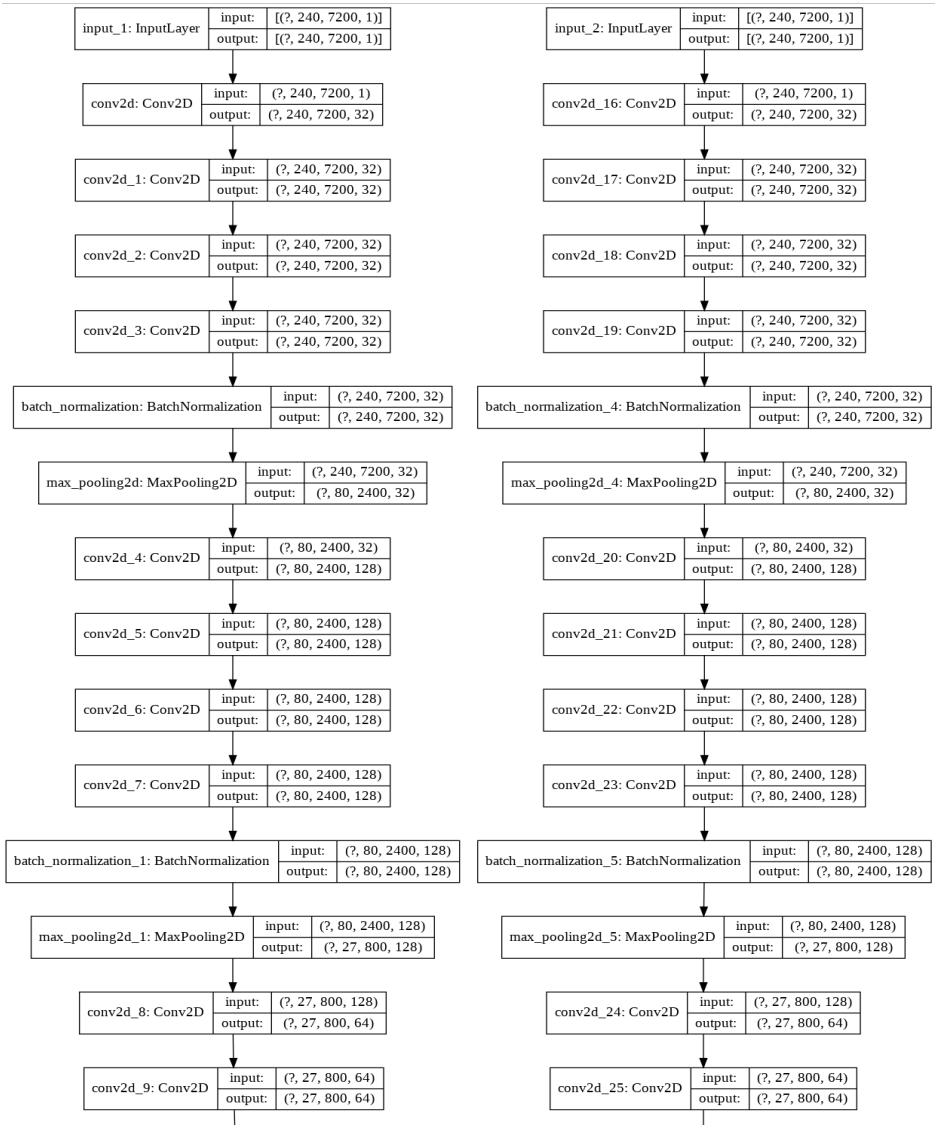


Figure C.3: % Change in Luminosity Between Months (Jan 2015-April 2020) With All Values Larger Than 2000 Set To Mean of Remaining Array.

Figure C.3 shows the percentage changes in Luminosity once the above methodology has been applied. This figure, when compared to figure 3.9, shows the scale of the fluctuations to be significantly (approximately 6000 times) smaller than when this processing has not been completed. The largest percentage changes are only between 0.4 % and -0.3%.

Appendix D

Appendix D - Neural Network Structures and Results



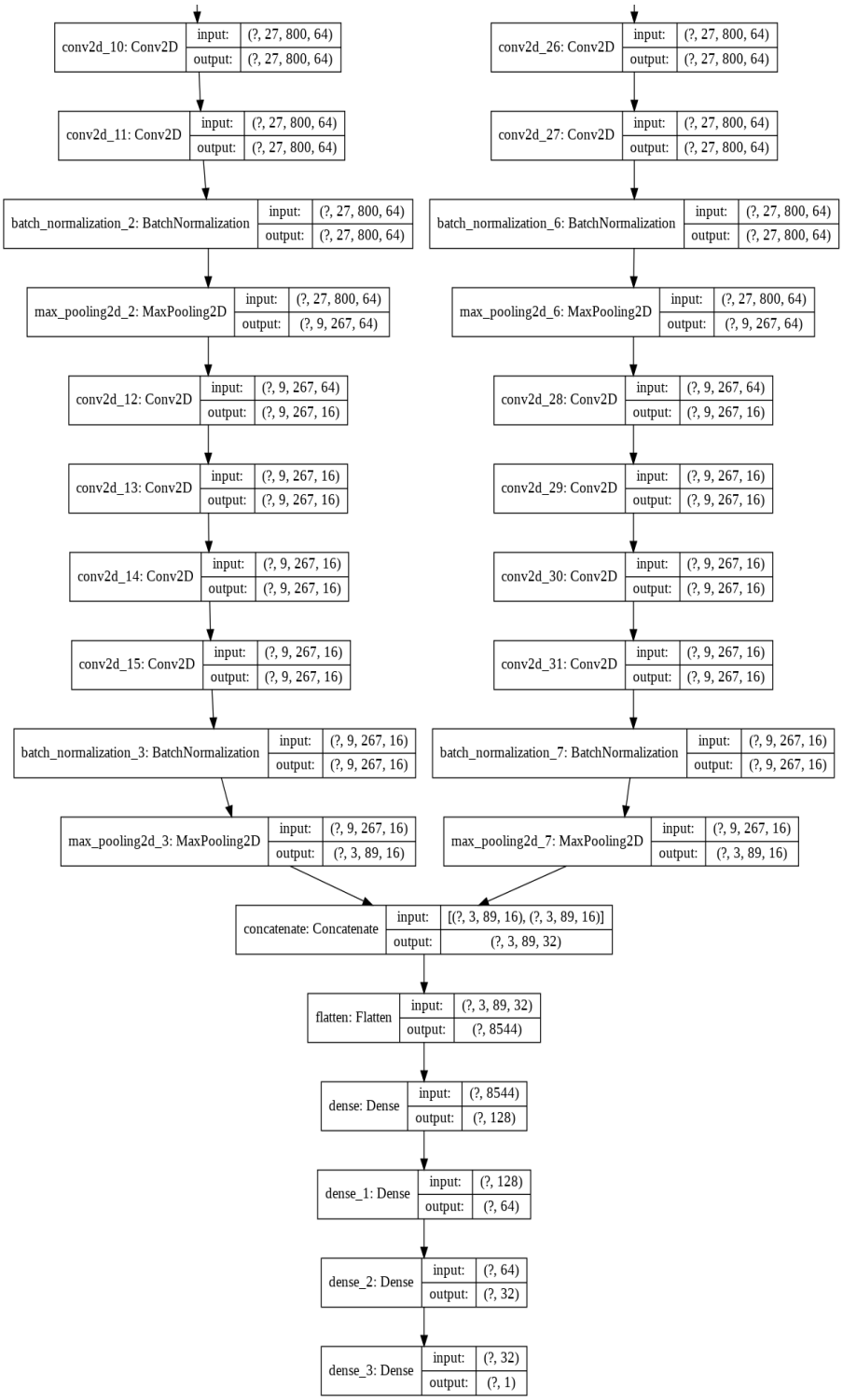
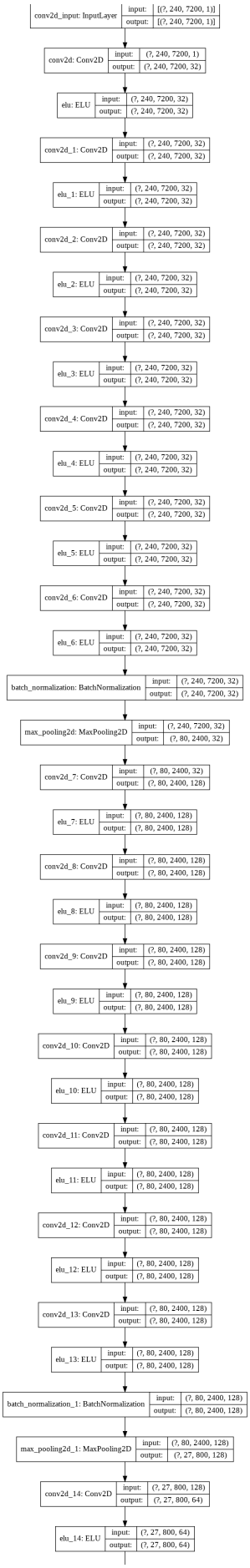


Figure D.1: CNN With Two Image Input Specification


conv2d_input: InputLayer | input: [(?, 240, 7200, 1)] | output: [(?, 240, 7200, 1)]
conv2d: Conv2D | input: (?, 240, 7200, 1) | output: (?, 240, 7200, 32)
elu: ELU | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
conv2d_1: Conv2D | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
elu_1: ELU | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
conv2d_2: Conv2D | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
elu_2: ELU | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
conv2d_3: Conv2D | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
elu_3: ELU | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
conv2d_4: Conv2D | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
elu_4: ELU | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
conv2d_5: Conv2D | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
elu_5: ELU | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
conv2d_6: Conv2D | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
elu_6: ELU | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
batch_normalization: BatchNormalization | input: (?, 240, 7200, 32) | output: (?, 240, 7200, 32)
max_pooling2d: MaxPooling2D | input: (?, 240, 7200, 32) | output: (?, 80, 2400, 32)
conv2d_7: Conv2D | input: (?, 80, 2400, 32) | output: (?, 80, 2400, 128)
elu_7: ELU | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
conv2d_8: Conv2D | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
elu_8: ELU | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
conv2d_9: Conv2D | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
elu_9: ELU | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
conv2d_10: Conv2D | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
elu_10: ELU | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
conv2d_11: Conv2D | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
elu_11: ELU | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
conv2d_12: Conv2D | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
elu_12: ELU | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
conv2d_13: Conv2D | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
elu_13: ELU | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
batch_normalization_1: BatchNormalization | input: (?, 80, 2400, 128) | output: (?, 80, 2400, 128)
max_pooling2d_1: MaxPooling2D | input: (?, 80, 2400, 128) | output: (?, 27, 800, 128)
conv2d_14: Conv2D | input: (?, 27, 800, 128) | output: (?, 27, 800, 64)
elu_14: ELU | input: (?, 27, 800, 64) | output: (?, 27, 800, 64)

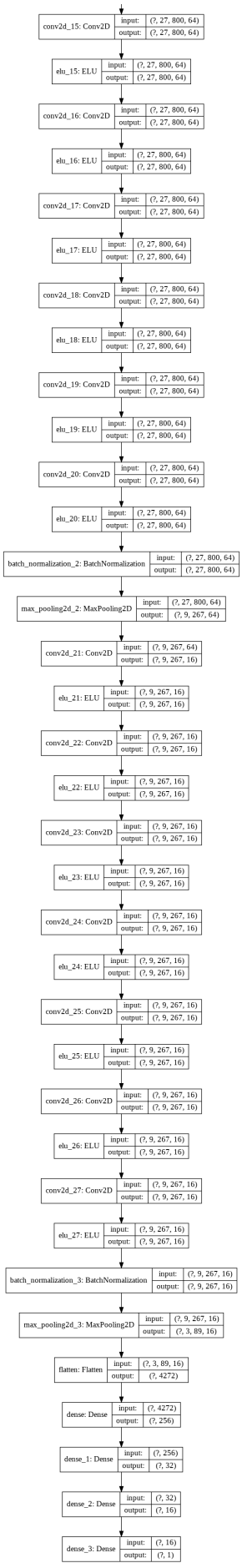


Figure D.2: CNN With One Image Input Specification

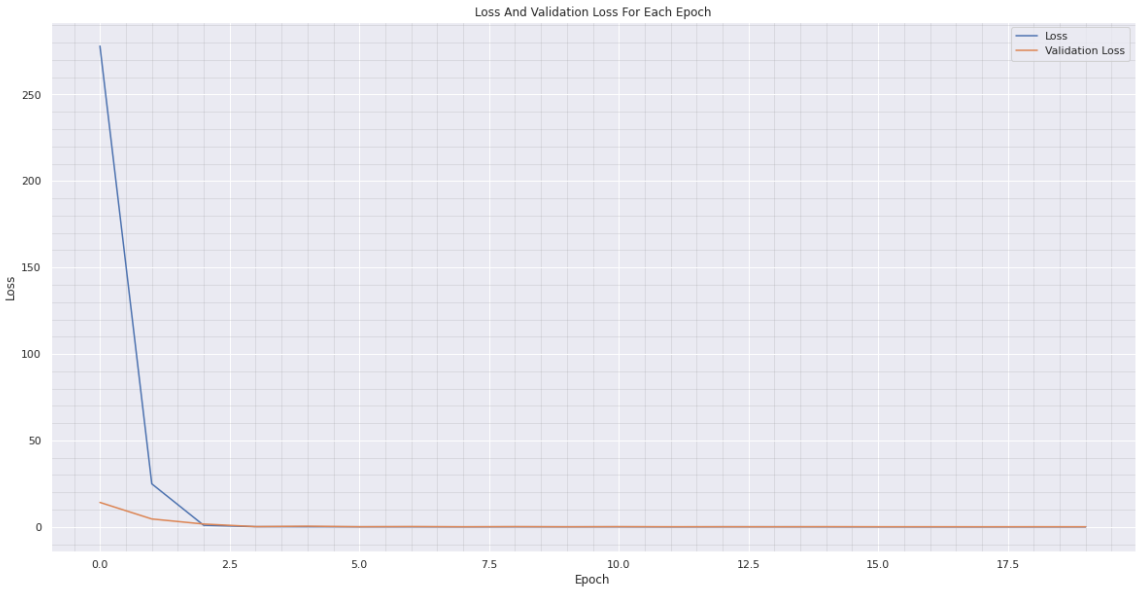


Figure D.3: Loss and Validation Loss For Model With Two Image Inputs