The Classification of Periodic Light Curves from non-survey optimized observational data through Automated Extraction of Phase-based Visual Features

Paul R. McWhirter
Computing and Mathematical Sciences
Liverpool John Moores University
Liverpool, UK.
P.R.McWhirter@2014.ljmu.ac.uk

Iain A. Steele
Astrophysics Research Institute
Liverpool John Moores University
Liverpool, UK.
I.A.Steele@ljmu.ac.uk

Dhiya Al-Jumeily
Computing and Mathematical Sciences
Liverpool John Moores University
Liverpool, UK.
D.Aljumeily@ljmu.ac.uk

Abir Hussain
Computing and Mathematical Sciences
Liverpool John Moores University
Liverpool, UK.
A.Hussain@ljmu.ac.uk

***Abstract*— We implement two hidden-layer feedforward networks to classify 3011 variable star light curves. These light curves are generated from a reduction of non-survey optimized observational images gathered by wide-field cameras mounted on the Liverpool Telescope. We extract 16 features found to be highly informative in previous studies but achieve only 19.82% accuracy on a 30% test set, 5.56% above a random model. Noise and sampling defects present in these light curves poison these features primarily by reducing our Periodogram period match rate to fewer than 5%. We propose using an automated visual feature extraction technique by transforming the phase-folded light curves into image based representations. This eliminates much of the noise and the missing phase data, due to sampling defects, should have a less destructive effect on these shape features as they still remain at least partially present. We produced a set of scaled images with pixels turned either on or off based on a threshold of data points in each pixel defined as at minimum one fifth of those of the most populated pixel for each light curve. Training on the same feedforward network, we achieve 29.13% accuracy, a 13.16% improvement over a random model and we also show this technique scales with an improvement to 33.51% accuracy by increasing the number of hidden layer neurons. We concede that this improvement is not yet sufficient to allow these light curves to be used for automated classification and in conclusion we discuss a new pipeline currently being developed that simultaneously incorporates period estimation and classification. This method is inspired by approximating the manual methods employed by astronomers.**

Keywords—Data analysis; Feedforward Networks; Light Curve Classification; Variable Stars; Visual feature extraction

I. INTRODUCTION

Time-domain Astronomy is an active field of discovery driven by recent technological advancements in observation, storage and data processing. Recent years have allowed for extended sky surveys such as the Sloan Digital Sky Survey (SDSS) [1] and the Gaia satellite which is mapping billions of stars allowing their fundamental parameters to be determined [2]. In the next few years this capability will grow through the deployment of even more powerful surveys such as the Large Synoptic Survey Telescope (LSST) generating approximately 20 TB of raw data every night [3]. These surveys are capable of regularly gathering data on wide regions of the sky. The regularity of this sampling is defined by the survey cadence. Each survey is optimized to a given cadence describing the approximate duration between observations of the same area of sky.

Despite the immense timescales involved in many Astrophysical processes, a number of variable phenomena occur within more human-comprehendible time scales. Few of these events are as well-studied as variable stars. These stars are at a volatile stage in their evolution resulting in perceived brightness fluctuations due to physical processes in their atmospheres. These stars allow for the study of stellar evolution and galactic structure [4]. Additionally, certain types of variable stars such as RR Lyraes or Delta Cepheids exhibit specific luminosity-period relationships that allow the determination of their distances from Earth [5, 6]. Other important variable light sources are eclipsing binary systems. These systems exhibit periodic brightness changes from binary stars eclipsing each other due to the orbital plane of the system having a low inclination relative to the Earth [7].

It is of great importance to reliably identify and monitor these objects and large wide-field sky surveys are an ideal method of accomplishing this task. The quantity of data from these surveys makes this a daunting exercise. Fortunately, the field of machine learning has provided techniques that can be

developed for the automated classification of light sources. There have been a number of studies investigating the production of both general purpose and more specific learned classification models through the extraction of useful features from the data of known variable objects.

Debosscher et al. developed a method of fitting harmonic models to light curves using multiple periods identified by a Lomb-Scargle Periodogram. These models were used to extract Fourier-based features for the production of learned classification models [8]. These features were extended into a set of 30 plus descriptive properties for light curves by adding non-periodic features by Richards et al [9]. These features were then processed into 16 highly informative features for general purpose variable star classification in the form of the Upsilon software package by Kim et al. [10]. Nun et al have collated these feature extraction methods into the Feature Analysis for Time Series package [11]. Pichara et al. have also proposed using meta-classification allowing the use of multiple high-performance, specific classification models named experts in general purpose classification tasks [12]. Puegert et al. have also proposed the extraction of shape-defining features from phase-folded light curves through the coefficients of fitted chains of polynomial models. These features were then used to classify eclipsing binary light curves using learned models built using a single hidden layer feedforward neural network [7].

These studies focused on datasets comprised of well-sampled fixed-cadence light curves. How would these methods perform on a different style of light curves derived from wide-field observations without a guaranteed cadence? These observations can have a significantly uneven distribution in time for individual light sources [13]. An example of one of these datasets is the observations produced by the Small Telescopes Installed at the Liverpool Telescope (STILT) [14]. The cameras are mounted to the frame of the Liverpool Telescope aiming co-parallel with the main telescope's field of view capturing a ten second exposure every minute whilst the telescope is in operation. They have no control over the position of the telescope and therefore no ability to influence their observational cadence. These observations from March 2009 to March 2012 have been processed into a dataset containing over 27 million individual light sources.

In this paper we introduce the initial results and problems from the application of the methods from these previous studies to the STILT observations and propose a method of automatically extracting shape-based features from the phase-folded light curves through the use of multiple neural network layers trained to recognize visual features mirroring the methodology employed manually by astronomers. In the future additional topologies will be introduced to further power this feature extraction and introduce knowledge of other phases through a phase-period similarity measure.

The rest of this paper is structured as follows. In Section 2, a selection of variable star light curves, generated from the STILT observations through coordinate comparison with the American Association of Variable Star Observers (AAVSO) variable star index catalogue is presented. This light curve dataset is then classified based on features proposed by previous studies. Section 3 introduces the proposed method for the extraction of visual features directly from a phase-folded representation of the light curves through the use of a two hidden layer feedforward neural network topology and shows early results from the application of this method to the same STILT dataset. In Section 4 the study is concluded and the currently-in-development proposed pipeline is described as an indication of the direction of future work.

II. STILT Variable Star Light Curves

The Small Telescopes Installed at the Liverpool Telescope (STILT) dataset is a wide field object SQL database. It contains 1.24 billion separate object observations of 27.74 million independent stellar objects. It was generated through the pre-processing of observational images gathered by the STILT instruments [14]. This database contains light curves for many objects, including many of unknown classification. Reliable class information is required for a subset of objects in the database in order to test classification methods on these light curves. The optimal method to extract this class information is through a comparison between the STILT data to a variable star catalogue. The American Association of Variable Star Observers (AAVSO) operates one of the largest and best-updated catalogues of nearby bright variable stars in the world, The AAVSO International Variable Star Index. This catalogue does not contain any of the AAVSO gathered light curves but it does contain data on 373,565 known variable stars including their name, coordinates in right ascension and declination and their currently identified class. The coordinates of these variable stars were matched to objects in the STILT database with a tolerance of 3.6 arc seconds (seemingly sufficient to avoid detection collisions between nearby stars) and a minimum of 100 individual observations. This resulted in the production of 12461 variable stars of various types. Ten variable star classes and super-classes were selected which were well represented in this dataset leaving 3011 corresponding objects. Table I demonstrates the class by class breakdown of this dataset.

TABLE I
3011 Object Dataset

Class	Dataset Statistics		
	Type	**Acronym**	**Count**
1	Delta Cepheid Variables	DCEP	132
2	Delta Scuti Variables	DSCT	499
3	Algol-Type Eclipsing Binaries	EA	684
4	Beta Lyrae Eclipsing Binaries	EB	242
5	W Ursae Majoris Eclipsing Binaries	EW	291
6	Mira Type Long Period Variables	M	149
7	Rotational Variables	ROT	492
8	RR Lyrae variables	RR	114
9	RS Canum Venaticorum Eclipsing Binaries	RS	192
10	Semi-regular Long Period Variables	SR	216

Following the selection of these 3011 variable light curves, the performance of the features used in previous high-performance general purpose classifiers was established. The 16 features used by Kim et al. [10] were chosen for this operation as they had been shown to be capable of reliably separating super-classes as well as achieving respectable inter-class accuracy. These features are shown in Table II.

TABLE II
KIM ET AL. 16 VARIABILITY FEATURES

Feature	Description	Reference
Period	Period derived by the Lomb-Scargle Periodogram	Kim et al. 2014
ψ^{η}	η of a phase-folded light curve	Kim et al. 2014
ψ^{CS}	Cumulative sum index of a phase-folded light curve	Kim et al. 2014
R_{21}	2^{nd} to 1^{st} amplitude ratio	Kim et al. 2014
R_{31}	3^{rd} to 1^{st} amplitude ratio	Kim et al. 2014
Φ_{21}	Difference between 2^{nd} and 1^{st} phase	Kim et al. 2014
Φ_{31}	Difference between 3^{rd} and 1^{st} phase	Kim et al. 2014
γ_1	Skewness	Kim et al. 2014
γ_2	Kurtosis	Kim et al. 2014
K	Stetson K index	Kim et al. 2014
Q_{3-1}	Difference between 3^{rd} and 1^{st} quartiles	Kim et al. 2014
A	A ratio of magnitudes brighter or fainter than the average	Kim et al. 2016
H_1	Amplitude from Fourier decomposition	Kim et al. 2016
W	Shapiro-Wilk normality test	Kim et al. 2016
m_{p10}	10^{th} percentile of slopes of a phase-folded light curve	Long et al. 2012
m_{p90}	90^{th} percentile of slopes of a phase-folded light curve	Long et al. 2012

The Lomb-Scargle Periodogram [15] utilised in this method operated over a linear frequency grid from the reciprocal of the total observation time of a light curve up to 20 cycles per day. The interval between candidate frequencies is shown in equation 1 where t_{max} and t_{min} are the last and first observation times respectfully.

$$f_{step} = \frac{0.25}{t_{max} - t_{min}} \quad (1)$$

The 16 features were plotted relative to the ten classes in this dataset. This was done to visually inspect how successful the features were at separating these classes in the feature space. Figure 1 clearly shows that these features have been extremely poor at accomplishing this with the STILT light curves. The ten classes are arranged from left to right in alphabetical order in this diagram.

On closer inspection, the Period feature appears to have a correct match rate of under 5% relative to the AAVSO period (which is treated as the ground truth in this study) for many of the classes. As the period, calculated by the Lomb-Scargle Periodogram, is the basis in which phase-folded light curves are generated, this inaccuracy heavily pollutes an additional 9 features. This is over half the number of features used in this analysis. As for the non-folded features, the distribution of these features amongst the classes appears to centre at or near their expected means. However, the range is much greater than expected increasing the overlap between classes. This is likely a result of the larger-than-usual noise threshold in the STILT data [13].

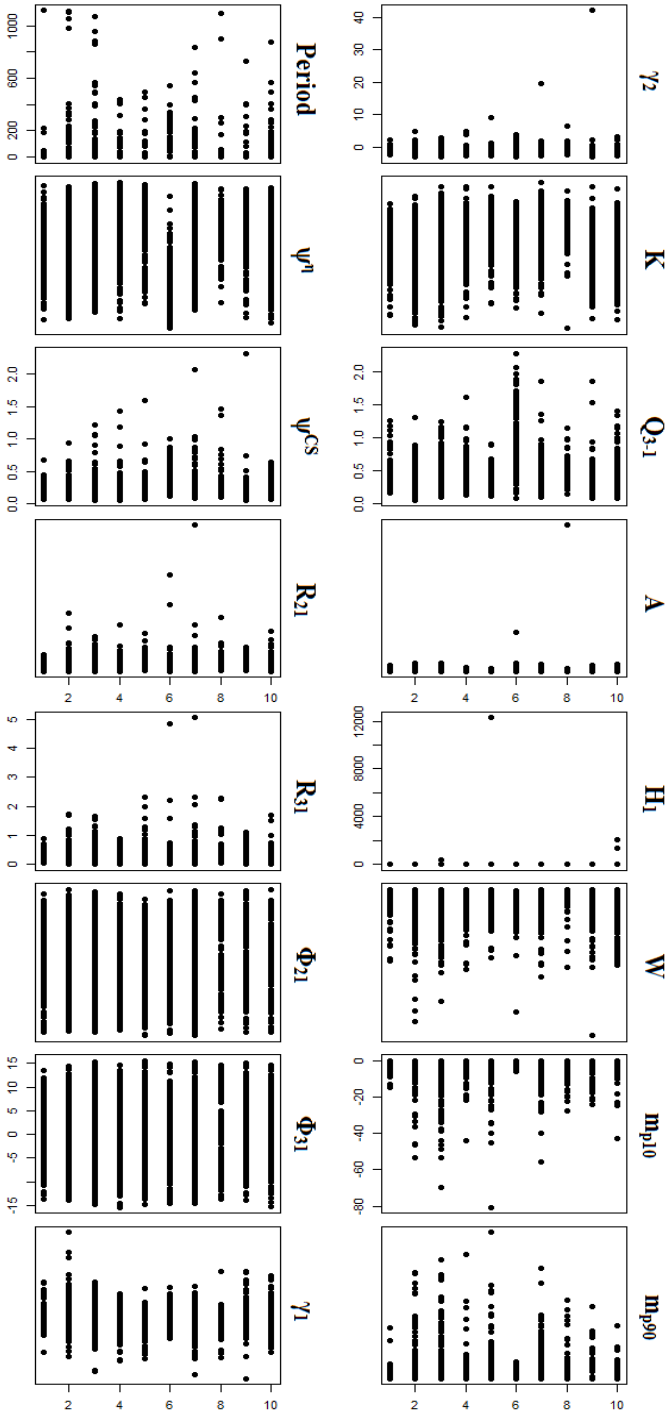

Fig. 1. Plot of each of the 16 features against the ten classes in alphabetical order. The features appear extremely poor at differentiating most classes with STILT light curves. This bodes extremely poorly on the performance of a learned classification model.

With these issues recognised and documented, we continued with these features in order to train a learned classifier. Previous studies make heavy use of the random forest algorithm yet in this study we will instead use a two hidden layer feedforward neural network. This is chosen as to directly compare these features to our proposed method in the next section.

The 3011 light curves are randomly sampled into a training set consisting of 2103 light curves and a test set consisting of 908 light curves. This is done whilst maintaining the ratio of each class in the subsets relative to the whole dataset. The training dataset was then used to train a two hidden layer feedforward neural network with 16 input neurons, 120 neurons in the first hidden layer, 80 neurons in the second hidden layer and 10 neurons in the output layer using a softmax classifier. Rectified Linear Units (ReLUs) are used for non-linearity and complexity control is introduced through a regularisation term valued at 10^{-3}. The network was trained using backpropagation over 50,000 iterations where an iteration is one forward and backward pass. Upon the completion of this task, the resultant model is used to predict the classes from the test dataset. The class with the highest predicted probability by the softmax function is assigned as the predicted class. A confusion matrix then compares the predicted classes to the actual classes and a set of specificities and sensitivities can be determined from this result. Figure 2 and Table III show the confusion matrix and the per-class specificity, sensitivity and balanced accuracy.

	Reference									
Prediction	1	2	3	4	5	6	7	8	9	10
1	11	2	6	3	0	2	3	1	1	2
2	5	37	47	16	8	2	24	2	12	11
3	17	68	97	31	45	8	63	26	25	30
4	0	0	0	0	1	0	2	1	1	0
5	1	4	5	2	5	0	2	0	1	1
6	3	5	5	3	1	28	5	1	2	9
7	3	25	42	14	22	0	45	3	10	4
8	0	1	2	1	3	1	1	0	1	1
9	0	3	1	0	2	0	1	1	2	1
10	0	5	1	3	0	4	2	0	3	6

Fig. 2. The confusion matrix for the trained model on the 16 Kim et al. features. There are many misclassifications and the classes with a higher population in the training data tend to be the preferred chose when the model cannot extract sufficient usable information.

TABLE IIII
ACCURACY OF THE FEEDFORWARD MODEL ON THE 16 KIM ET AL. FEATURES

Class	Sensitivity	Specificity	Accuracy
DCEP	0.2750000	0.9769300	0.6259700
DSCT	0.2466700	0.8322300	0.5394500
EA	0.4709000	0.5535000	0.5122000
EB	0.0000000	0.9940050	0.4970020
EW	0.0574710	0.9804880	0.5189800
M	0.6222200	0.9605600	0.7913900
ROT	0.3040500	0.8379400	0.5710000
RR	0.0000000	0.9873900	0.4936900
RS	0.0344830	0.9893990	0.5119410
SR	0.0923080	0.9786220	0.5354650

These results are obviously very poor with only the Delta Cepheid and Mira classes climbing to even a respectable classification rate. The overall accuracy of this model was 25.47%. To compare the performance of this model relative to a random prediction, we also train a 'random model' as described by Kim et al. [10]. This is accomplished by randomising the class labels and retraining the network in this new configuration. As the expected feature ranges of the classes no longer lay together, this model should approximate random prediction. Table IV show the statistics of this random model compared to the previously trained model and can be directly compared to Table III. The random model has an overall accuracy of 19.82% and therefore the 16 features are only a 5.56% accuracy improvement over random selection. Due to the random sorting, the Delta Cepheid and Mira classes no longer achieve higher results and have descended to the level of the other classes. As expected, these features have been too heavily poisoned by noise and errors to be of any use in this classification task. Additionally, as was seen in the real classification, the almost complete loss of information resulted in many objects being placed in the class with the most populous training data. This can be augmented by altering the probability boundary in which a light curve will be assigned a given class however; this is not performed during this study.

TABLE IV
ACCURACY OF THE FEEDFORWARD MODEL ON THE RANDOMISED DATA

Class	Sensitivity	Specificity	Accuracy
DCEP	0.0256410	0.9953970	0.5105190
DSCT	0.1842100	0.8201100	0.5021600
EA	0.6000000	0.4011000	0.5006000
EB	0.0000000	0.9939170	0.4969590
EW	0.0111110	0.9853300	0.4982210
M	0.0000000	0.9976770	0.4988390
ROT	0.1862100	0.8099600	0.4780800
RR	0.0000000	0.9988710	0.4994360
RS	0.0000000	0.9976880	0.4988440
SR	0.0000000	1.0000000	0.5000000

By using the probabilities predicted by the softmax function for each of the light curves across the two models, two Receiver Operating Characteristic (ROC) Curves are also plotted. These curves are a measure of a class's true positive rate against the false positive rate with the ideal classifier maximising the true positive rate whilst minimising the false positive rate. Therefore, the better performing a class, the closer it will deviate towards the top left corner from the random-state as a straight line with slope 1 shown by the dotted black line in the figures. Figure 3 shows the ROC curve generated by the 16 feature model and figure 4 shows the ROC curve produced by the random model associated with the 16 features. Each line is related to a one-vs-many prediction on a specific class given by the line colour in the legend. This is performed by assigning a class label of 1 to the appropriate class and a label of 0 to all other classes. Figure 3 clearly shows a number of classes having accuracy greater than random, but not by an impressive level due to the poor features whereas the random model in figure 4 has every class grouped on the dashed line.

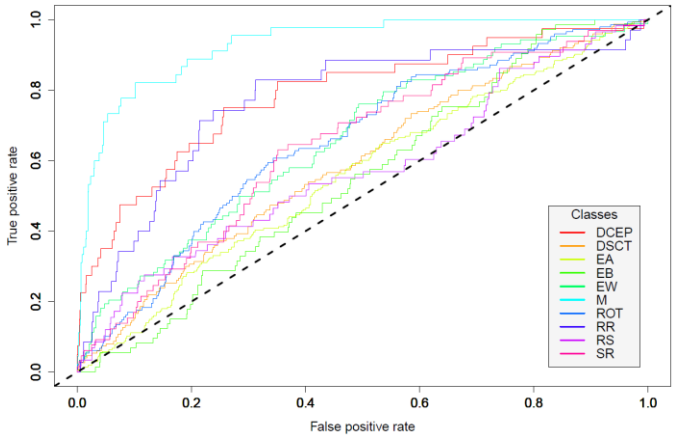

Fig. 3. ROC curve for the 16 feature trained on the test set. The Mira, Delta Cepheid and Rotationally spotted star classes appear to have the most success with many others unacceptably close to the dashed random line yet still on the correct side of it.

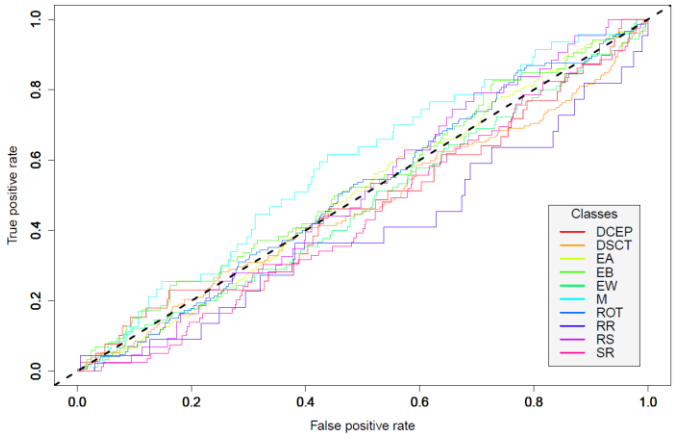

Fig. 4. ROC curve for the 16 feature random model on the test set. As expected, the classes all cluster around the dashed random line.

III. Visual Feature Extraction

These 16 features have been shown to be fully capable of training useful classifiers yet on this dataset they have almost completely failed to do so. To understand why this occurred, it is important to understand why the features are important. In essence, they are attempting to describe the shape of brightness changes through the dominant periodic variation by 'folding' all the gathered data points into one waveform. This is very useful in astronomy due to the limitations in gathering data. In fact, this is one of the most powerful techniques in eliminating sampling issues as long as the light curve does have a dominant period. For non-periodic variable objects in astronomy, such as transient light sources, other approaches must be considered. In the case of many periodic variable objects, the shapes of the light curves in these phase-folded representations carry significant information about the class of the light source. Figure 5 shows an example of three of the light curves in this dataset, a Mira-type Long Period Variable, an Algol-type eclipsing binary and a Delta Cepheid. These light curves have been folded at the AAVSO period of the associated objects. Therefore, a light curve must clearly demonstrate these shape features in order for the 16 Kim et al. features (and many more) to extract enough of this information from any noise.

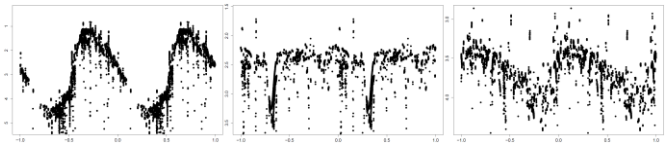

Fig. 5. STILT dataset folded Light curves of the star Mira (Mira class) with a 332 day period, Algol (Algol-type eclipsing binary) with a 2.86 day period and Eta Aquilae (Delta Cepheid class) with a period of 7.18 days. The shape of each light curve is distinctive to the associated class.

The light curves in figure 5 are fairly typical of the better sampled light curves from the STILT database yet they clearly show some clues into the apparent issues. Firstly, there are a lot of points with significant noise. This is possibly instrumental in nature but is much more likely due to a number of simplifications made to reduce the computational load of the pre-processing pipeline. This noise is likely to be the cause of both the larger range on the non-periodic features as well as the cause of the very low period match rate from the Lomb-Scargle Periodogram. Secondly, whilst the examples in figure 5 are well sampled across the whole phase space, there are other light curves that lack this due to the highly variable cadence of the STILT observations. This means that important shape features may only be partly present and not to the level required for the extracted features in previous studies.

Yet, despite these obvious limitations, human astronomers can still look at these light curves and recognise the main shape patterns. Therefore it seems reasonable to conclude that even in the more poorly sampled, noisy STILT light curves, there are still features that have not yet been extracted which are being gathered for manual classification. Ideally the models used to fit the light curves should attempt to parameterise the shape of the actual variable object classes rather than some predefined or abstract form. This can be done by determining the specific form of different astrophysical signals directly from the astrophysics driving the variability of these object types. This is quite an undertaking further complicated by a lack of consensus about the dominant physical processes shaping these variabilities in many classes. Therefore, it would seem to be more appropriate to have a model that can identify the patterns in the shape of the light curves without requiring an underlying physically produced model. This can be accomplished through a learning process applied to visualized examples of light curves. Over the last decade, neural networks have been developed into platforms for visual reasoning [16]. The ImageNet classification is a good example, a large dataset of images collected into 1000 classes. Respectable classification accuracy has been found through the use of deep networks with convolutional layers for visual feature extraction [16]. We attempt to replicate these visual feature extraction layers through the construction of hidden layers tuned to find visual features. As this is just the initial investigation, convolutional layers have not yet been utilised and this does result in limitations discussed shortly.

We first phase-folded the light curves for each of the STILT dataset light sources. This task required a candidate period. As the Lomb-Scargle Periodogram is performing poorly on our data, we instead used the AAVSO period for the objects. We concede that this gave a lot of important information freely to this model straining the comparison we hoped to make with the 16 feature model. Despite this fact we

continued with this approach. In the conclusion we discuss this aspect of the study further as well as proposing a possible solution to the period matching problem.

The phase-folded light curves were used to generate pixelated images. This has a number of important uses. First we can guarantee an identical number of inputs into our neural network regardless of the sampling of the light curve. Second, it can be minimized to a level which optimizes for computational cost whilst simultaneously eliminating much of the noise present in the light curves. For this task we decided to transform the light curves into 28x28 pixel images giving 784 input pixels. There were also two hyper parameters in this operation. A sampling level which indicates how well populated an individual pixel must be to turn on relative to the most populated pixel in the light curve. The other hyper parameter is whether pixels that turn on have an analogue gradient based on their data point population or alternatively they are just fully on if any data point is present above the sampling threshold. The activations were normalised so the most populated pixel of any light curve is always on at the maximum value. In this study the sampling level was set at a minimum of one fifth of the value of the most populated pixel to turn on and that the pixels are just binary on or off with no analogue value.

Each light curve produced four magnitude-scaled images of height 3, 4, 5 and 6 times the standard deviation of the light curve centred on its weighted mean. Like the previous 16 feature models, the dataset was split into a 70%/30% training set and testing set. The same feedforward network topology was used with two hidden layers, the first with 120 neurons and the second with 80 neurons, regularization for complexity control and ReLUs for non-linearity. The primary difference from the first models was there were now 784 input units where each one is the value of a specific pixel from a concatenated 28x28 image representation vector, -0.5 for an off (black) pixel and +0.5 for an on (white) pixel. This model was also trained for 50,000 iterations. The resulting model was then used to predict the class labels for the testing dataset based on their image representations. This produced the confusion matrix statistics shown in Table V. This model was the first trained using this method and is also the first to not classify unsure objects into the most populous class, instead choosing the class that the light curves image most closely matched. The model has an overall accuracy of 29.13% and this model showed improved performance on the Delta Scuti stars.

TABLE V
ACCURACY OF THE FEEDFORWARD MODEL ON THE IMAGE MODEL

Class	Sensitivity	Specificity	Accuracy
DCEP	0.2500000	0.9737900	0.6119000
DSCT	0.5466700	0.8984200	0.7225400
EA	0.1881100	0.8920900	0.5401000
EB	0.0582190	0.9176650	0.4879420
EW	0.3380700	0.8783500	0.6082100
M	0.6333300	0.9690000	0.8011700
ROT	0.3581100	0.7710500	0.5645800
RR	0.1928570	0.9630580	0.5779580
RS	0.0732760	0.9520590	0.5126670
SR	0.1115380	0.9694540	0.5404960

A random model was also trained on the image based representations using the same method discussed above. Table VI presents the confusion matrix statistics of this random model. Again, the predictions tend to place the test light curves into the most populous training classes and the balanced accuracy of most classes is close to 50% random. However, it is of interest to note that the overall accuracy is much lower at 15.97% meaning that the actual model has a 13.16% improvement over the random model.

TABLE VI
ACCURACY OF THE FEEDFORWARD MODEL ON THE IMAGE RANDOM MODEL

Class	Sensitivity	Specificity	Accuracy
DCEP	0.0000000	1.0000000	0.5000000
DSCT	0.1019100	0.9081200	0.5050200
EA	0.2229100	0.7809000	0.5024100
EB	0.0000000	0.9994035	0.4997018
EW	0.1370300	0.8312600	0.4841400
M	0.0000000	1.0000000	0.5000000
ROT	0.4571400	0.5459300	0.5015400
RR	0.0000000	1.0000000	0.5000000
RS	0.0000000	1.0000000	0.5000000
SR	0.0689660	0.9338480	0.5014070

ROC curves were also plotted for these two models. Figure 6 shows the curves from the image representation model and figure 7 demonstrates the curves from the random image representation model. Whilst the data is still not resulting in acceptable classification accuracy across all classes, the models do show that features automatically extracted by neural network layers trained to recognise visual shapes can be used in a classification task.

This network is also extremely limited in the visual features it can extract. For example, despite attempts to position certain magnitude features at specific phases, noise quite often causes these features to be placed at slightly different phases. This results in the requirement of any visual feature layer to implement translation invariance. This can be accomplished by neural networks using convolutional layers [16] but this has not been implemented in these models, which is a big limitation.

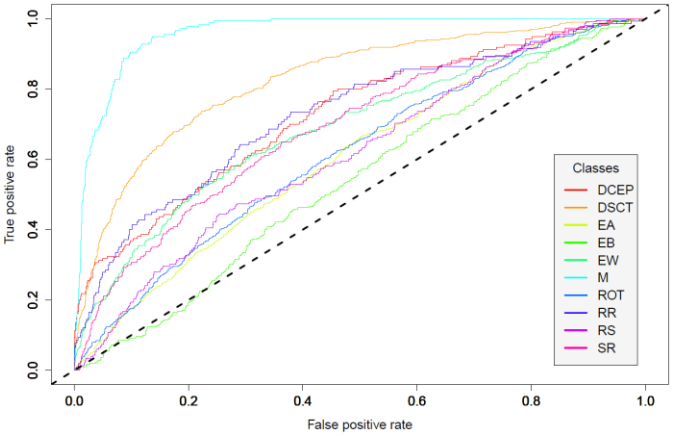

Fig. 6. ROC curve for the image representation model on the test set. The Delta Cepheid class is weaker than in the 16 feature model. Mira class stars remain strong and surprisingly the normally noisy Delta Scuti stars have improved accuracy in this model.

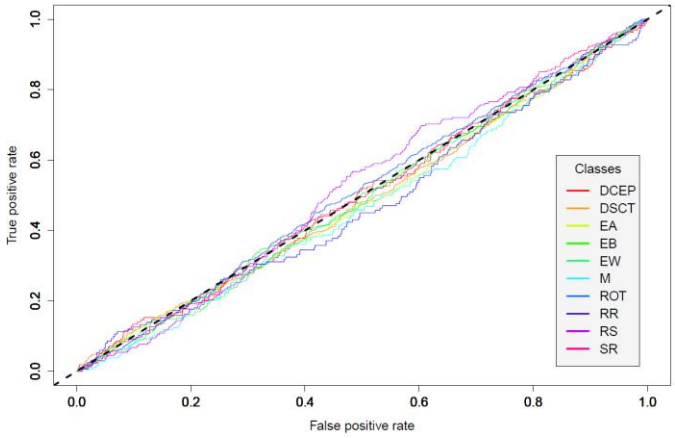

Fig. 7. ROC curve for the random image representation model on the test set. As expected, the classes all cluster around the dashed random line.

Finally, to show that this topology can be scaled up to improve feature extraction, a new model was trained using a similar method but now the two hidden layers had 1200 and 800 neurons, a substantial increase in capability. The resulting model improved its overall accuracy from 29.13% to 33.51% demonstrating the scalability of this method. It is likely this improvement was due to increased offsets from the current disadvantage of no translational invariance. The ROC curves generated from this large model are shown in figure 8.

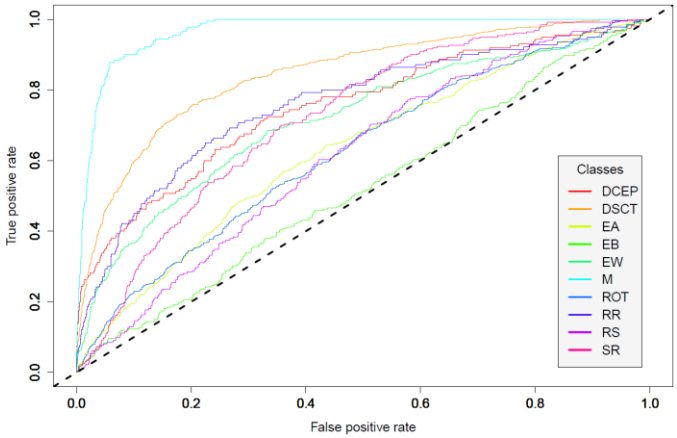

Fig. 8. ROC curve for the larger image representation model on the test set. Almost every class displays a small amount of improvement from this larger model due to the increased feature extraction capability of the network.

The first layer of the neural network is stored in a weight matrix of size 784xH where H is the number of neurons in the hidden layer. This can be unravelled into H 28x28 neurons and plotted as 28x28 images. Each image indicates the activation of that specific neuron based on the 28x28 input images. Therefore by plotting these out it is possible to visually inspect the trained neurons and recognise specific features they have trained to activate or not activate upon. Figure 9 shows the first 44 neurons learned by the normal-size image representation model. Of these neurons it is clear they are trained to recognise the variations within the light curves across the phase. Neuron 35 is of specific interest as it appears to be trained to activate upon the detection of a sinusoidal signal yet has a Delta Cepheid style dark area across it indicating that the saw-tooth shaped light curves of this class would supress the activation of this specific neuron.

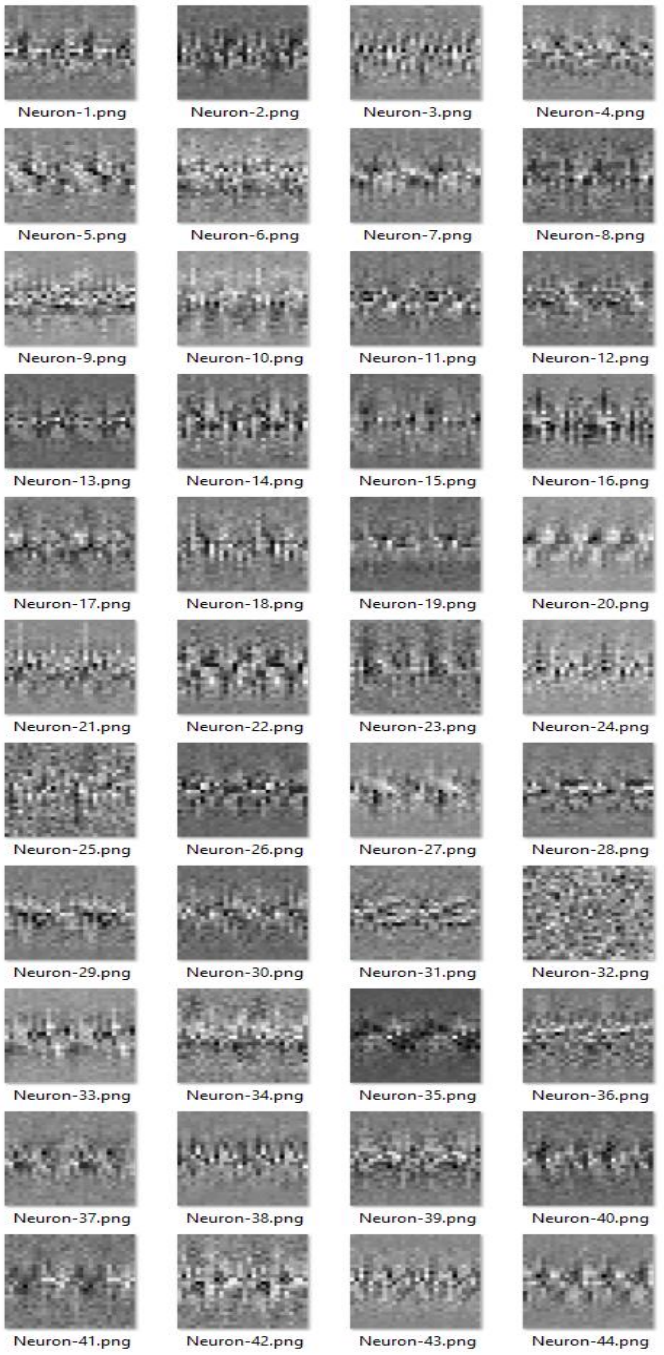

Fig. 9. The first 44 neurons learned by the 120 neuron first hidden layer in the image representation model. The neurons clearly show shape features resulting in both activations and deactivations of specific neurons.

IV. Conclusions and Further Work

The visual identification of folded light curves, whilst seemingly struggling with the straight classification of the STILT light curves, can also be considered the basis for a new approach to period estimation and classification. A major weakness in many Periodograms when applied to the STILT data is the periodic model utilised for fitting the light curve data by these algorithms. These models can be too simple such as the Lomb-Scargle Periodogram fitting a non-harmonic

sinusoidal model to the light curves [15]. Many variable objects have light curve profiles that are at best vaguely sinusoidal. On the other end of the spectrum, the non-parametric models such as the String Length Lafler Kinman technique are also vulnerable as they do not make any assumption on the shape of the light curve [17].

Taking a step back, it seems logical to ask the question: How does an astronomer manually decide the classification of a newly discovered variable object? This process usually involves running a Periodogram on the light curve as the 16-feature method has done followed by manually folding the light curve around a number of the candidate periods that are not known spurious periods. From these folded light curves, the astronomer will then make a decision to which period the resulting folded light curve looks similar to a known variable object class. Occasionally, the folded light curves will show that the candidate period is a multiple, a submultiple or a close miss of a period that correctly folds the light curve to a known object class. This manual method involves a lot of visual reasoning on the behalf of the astronomer but often the astronomer will produce an accurate result although it is occasionally biased based on the astronomers personal experience.

It would seem like a reasonable idea to have an automated algorithm attempt to replicate this method but with enough experience of the whole problem domain to adequately eliminate any bias. This is the basis of this proposed pipeline, teach computers to automate classification and period estimation tasks by performing visual reasoning on folded light curves like a human astronomer would. Instead of relying on an improved Periodogram, the pipeline will simultaneously deal with the period estimation and classification tasks through the use of the folded light curve transformation. A training phase would be required to produce multiple class-specific learned models for classification, not of a given light source's type, but rather whether the current candidate period is correct, a multiple, a near miss or a complete miss for a given object class. Input for training would be in the form of a set of light curves where each light curve is represented as a set of folded light curves based on a number of strong peaks from an algorithm such as the Lomb-Scargle Periodogram, along with a ground truth class label and the true astrophysical period. The light curves are folded and then used to generate models using a visual representation. These models can then be used to estimate the period and class of unknown light curves. First, a periodogram produces a subset of candidate frequencies from an initial frequency spectrum for an unknown light curve. The light curve is then folded multiple times for each of these candidate frequencies. Each candidate-folded light curve is classified into how far it deviates from the correct period, identifying multiples, submultiples and near misses. It is expected that when folded at or near the correct astrophysical period, the shape of the folded light curve will appear similar to the learned representations of the associated class.

ACKNOWLEDGMENT

The Liverpool Telescope is operated on the island of La Palma by Liverpool John Moores University in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofisica de Canarias with financial support from the UK Science and Technology Facilities Council [18]. We acknowledge with thanks the variable star observations from the AAVSO International Database contributed by observers worldwide and used in this research.

REFERENCES

[1] D. G. York, J. Adelman, et al., "The Sloan Digital Sky Survey: Technical Summary," The Astronomical Journal, vol. 120, 3, pp. 1579–2000, 2000.

[2] M. A. C. Perryman, K. S. de Boer, G. Gilmore, et al., "GAIA: Composition, formation and evolution of the Galaxy," Astronomy and Astrophysics, vol. 369, pp. 339–363, 2001.

[3] Z. Ivezic, J. A. Tyson, B. Abel, et al., "LSST: from science drivers to reference design and anticipated data products," ArXiv e-prints [arXiv:0805.2366], 2008.

[4] L. Eyer, N. Mowlavi, "Variable stars across the observational hr diagram," Journal of Physics: Conference Series, vol. 118, 1, 012010, 2008.

[5] A. K. Vivas, R. Zinn, P. Andrews, et al., "The QUEST RR Lyrae Survey: Confirmation of the Clump at 50 kiloparcsecs and Other Overdensities in the Outer Halo," The Astrophysical Journal, vol. 554, 1, pp. L33–L36, 2001.

[6] T. R. Bedding, A. A. Zulstra, "HIPPARCOS Period-Luminosity relations for Mira and semiregular variables," The Astrophysical Journal, vol. 506, pp. 47–50, 1998.

[7] M. Paegert, K. G. Stassun and D. M. Burger, "The EB Factory Project. I. A Fast, Neural-net-based, General Purpose Light Curve Classifier Optimized for Eclipsing Binaries," The Astronomical Journal, vol. 148, 2, article id. 31, 16, 2014.

[8] J. Debosscher, L. M. Sarro, et al., "Automated supervised classification of variable stars I. Methodology," Astronomy and Astrophysics, vol. 475, pp. 1159–1183, 2007.

[9] J. W. Richards, D. L. Starr, et al., "On Machine-Learned Classification of Variable Stars with Sparse and Noisy Time-Series Data," The Astrophysics Journal, vol. 733, 1, pp. 10–32, 2011.

[10] D.-W. Kim and C. A. L. Bailer-Jones, "A package for the automated classification of periodic variable stars," Astronomy and Astrophysics, vol. 587, A18, 2016.

[11] I. Nun, P. Protopapas, et al., "FATS: Feature Analysis for Time Series," ArXiv e-prints [arXiv:1506.00010], 2015.

[12] K. Pichara, P. Protopapas and D. León, "Meta-Classification for Variable Stars," The Astrophysical Journal, vol. 819, 1, 2016.

[13] P. R. McWhirter, S. Wright, I. A. Steele, D. Al-Jumeily, A. Hussain, P. Fergus, "A Dynamic, Modular Intelligent-Agent Framework for Astronomical Light Curve Analysis and Classification," Lecture Notes in Computer Science, vol. 9771, pp. 820–831, 2016.

[14] N. R. Mawson, I. A. Steele and R. J. Smith, "STILT: System design and performance," Astronomische Nachrichten, vol. 334, 7, pp. 729–737, 2013.

[15] J. D. Scargle, "Studies in Astronomical Time Series Analysis. II. Statistical aspects of spectral analysis of unevenly spaced data," The Astrophysical Journal, vol. 263, pp. 835–853, 1982.

[16] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems 25 (NIPS 2012), 2012.

[17] D. Clarke, "String/Rope length methods using the Lafler-Kinman statistic," Astronomy and Astrophysics, vol. 2, no. 386, pp. 763–774, 2002.

[18] I. A. Steele, R. J. Smith, et al., "The Liverpool Telescope: performance and first results," Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 2004.