Recycling Price Prediction of Renewable Resources

Ye Lu
University of Electronic Science and Technology of China
Chengdu Sichuan China
luy2333yyy@gmail.com

XinLei Chen
Carnegie Mellon University
Pittsburgh PA USA
xinlei.chen@sv.cmu.edu

BoJie Wang
HuanJia Group
bojiewang@hotmail.com

TengYue Wang
HuanJia Group
wangtengyuewty@163.com

Pei Zhang
Carnegie Mellon University
Pittsburgh PA USA
peizhang@cmu.edu

Yong Li
Tsinghua University
Beijing China
liyong07@tsinghua.edu.cn

ABSTRACT

In the renewable resource recycling market, the recycling price is a key factor which can influence the recycling market. The price prediction of renewable resources is important, which is helpful to guide the development of the market. However, it is difficult to make accurate predictions because the data of recycling prices is highly random and complex. Moreover, there is a time lag between the predicted prices and the accurate prices. In this paper, we propose a combined model to solve these problems. Our model decomposes the prediction into two parts: trend price prediction with Moving Average (MA) and residual price prediction with Empirical Mode Decomposition (EMD) and Long Short-Term Memory (LSTM) neural network. Evaluations on a real-world dataset show that our model outperforms those classical prediction models with the error reduced by over 70% and solves the time lag problem.

CCS CONCEPTS

• Information systems → Data mining; • Applied computing → Computers in other domains;

KEYWORDS

Renewable Resource, Time Series, Price Prediction, Moving Average, Empirical Mode Decomposition, Long Short-Term Memory

UbiComp/ISWC '19 Adjunct, September 9–13, 2019, London, United Kingdom

ACM ISBN 978-1-4503-6869-8/19/09…$15.00
https://doi.org/10.1145/3341162.3349326

ACM Reference format:

Ye Lu, XinLei Chen, BoJie Wang, et al. 2019. Recycling Price Prediction of Renewable Resources. Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the 2019 International Symposium on Wearable Computers *(UbiComp/ISWC '19 Adjunct). ACM, London, United Kingdom, 6 pages.* https://doi.org/10.1145/3341162.3349326

1 Introduction

The industry of renewable resources recycling plays an important role in the economic development. In 2018, the total amount of major renewable resources recycled in China reached 302 million tons and about 974.5 billion yuan. And in the recycling industry, recycling price is one key factor influencing its development. For example, it could bring effects to the rate of recycling[1] or sway the relevant decisions of the government[2]. Thus, it is a meaningful task to make predictions for daily recycling price of renewable resources, which is helpful to guide the development of the recycling market.

Typically, the recycling price data is in the form of time series, which is a series of data points listed in time order. The form of time series is common in economic data and there are a lot of researches on time series prediction. They can be roughly divided into model based methods and data driven methods. In detail, model based methods are represented by the Autoregressive Integrated Moving Average (ARIMA) model, which is based on Statistics. ARIMA model applies to the prediction of linear time series and has a good effect in many aspects[3], [4]. Data driven methods are mainly based on machine learning and deep learning. Support-Vector Regression (SVR) model is one of typical time series prediction models in machine learning. It is good at dealing with nonlinear problems with small data size. [5] carried out their work based on SVR model. For another, many experiments have proved that deep learning methods are quite effective for time series prediction. Especially, LSTM network has a better

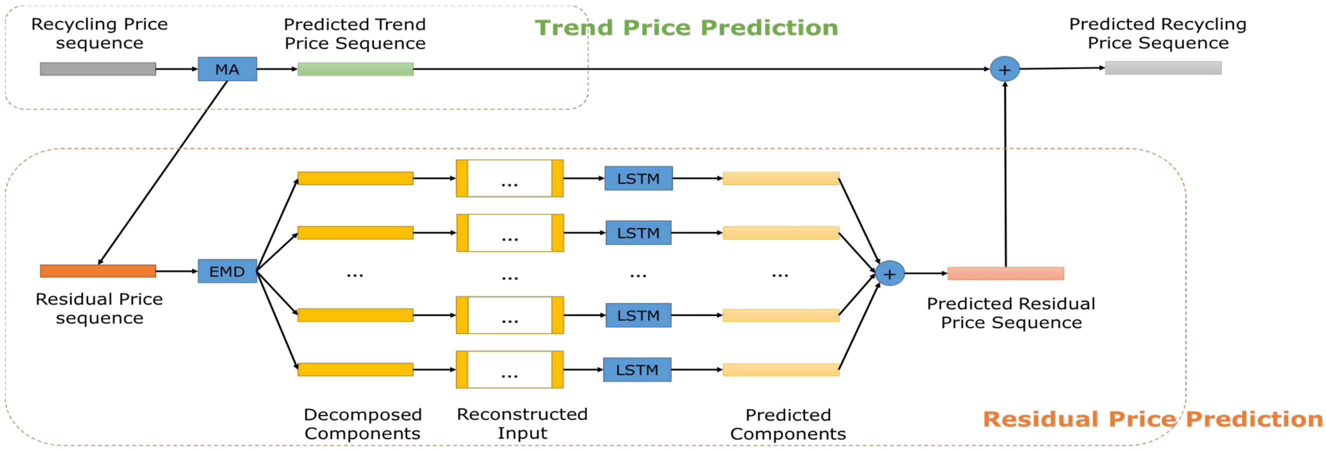

Figure 1: Structure of MA-EMD-LSTM Model. The green part is the trend price prediction for the long-term trend. The orange part is the residual price prediction for the short-term volatility.

performance and is popular in the study of nonlinear and unsteady time series prediction[6], [7].

In spite of so much research work on time series prediction, there is little relevant work on the prediction for the recycling price of renewable resources due to two reasons. For one thing, the recycling price series tends to be unsteady and nonlinear, which may have a long-term trend and short-term volatility simultaneously. For another, there is a strong self-correlation in the target time series. Using a classical model to make recycling price prediction has a problem of time lag, which fails to predict the small changes between two successive days.

In order to solve these problems, this paper presents MA-EMD-LSTM, a combined model making recycling price prediction with high accuracy. Since the changes of the recycling price series are complex, our model divides the task into two parts, trend price prediction for the long-term trend and residual price prediction for the short-term volatility. In the trend price prediction, MA model is used to predict the trend values and obtain the corresponding residual sequence. And the residual price prediction is the focus of our model, which aims to predict the small changes of daily recycling prices. It could help solve the problem of time lag in the predicted results. To be specific, EMD method is used to process the data first. It decomposes the residual sequence into a series of more stable component sequences. And then LSTM networks are used to make prediction for them separately. Using EMD method to preprocess the input data of the LSTM networks help get a better performance than using a single LSTM network. Experiments on a real-world dataset show that our model outperforms those classical prediction models, with the error reduced by over 70%.

The main contributions of this paper are summarized as:

1. We decompose the task of recycling price prediction into long-term trend price and short-term residual price prediction based on the time series characteristics of the recycling price.
2. We propose a combined model MA-EMD-LSTM, in which MA and LSTM deal with long-term and short-term price prediction respectively, while EMD decomposes the short-term residual price before LSTM for better performance.
3. We evaluate our MA-EMD-LSTM model with a real world dataset, which demonstrates the effectiveness of our model.

2 Problem Definition

In the task of recycling price prediction, the history data of one resource's recycling price is given, which is in the form of time series. Our work is to make prediction for the next day's recycling price based on the given history data. In detail, we denote a recycling price series as $S = x_1, x_2, \cdots, x_t$ and a subsequence of the last l time steps as $S_{t-l+1,t} = x_{t-l+1}, x_{t-l+2}, \cdots, x_t$. Then the task is to build a model to get a prediction function for the next day's price based on the last l time steps of data: $\hat{x}_{t+1} = f(S_{t-l+1,t}) = f(x_{t-l+1}, x_{t-l+2}, \cdots, x_t)$. We aim to minimize the error between the predicted price $\hat{x}_{t+1}$ and the actual price x_{t+1} by design an appropriate model. Our target of the task in this paper is as follows:

$$\min L = \hat{x}_{t+1} - x_{t+1}$$
$$\text{s.t. } \hat{x}_{t+1} = f(S_{t-l+1,t}) = f(x_{t-l+1}, x_{t-l+2}, \cdots, x_t) \quad (1)$$

3 MA-EMD-LSTM Model

In this section, we will introduce the details of our proposed model MA-EMD-LSTM. The structure of our model is shown in Figure 1.

According to our dataset, the recycling price series is usually unsteady, aperiodic and nonlinear, having the long-term trend and short-term volatility at the same time. Those classical models for time series prediction failed to reach a high prediction accuracy in this task. For they are good at dealing with linear and stationary series and have a problem of time lag for recycling price data. Figure 2 shows the result of a classical model, illustrating the problem of time lag. Therefore, we propose a combined model MA-EMD-LSTM to solve these problems. In detail, our model has two parts, MA for long-term trend price prediction and EMD-LSTM for short-term residual price prediction.

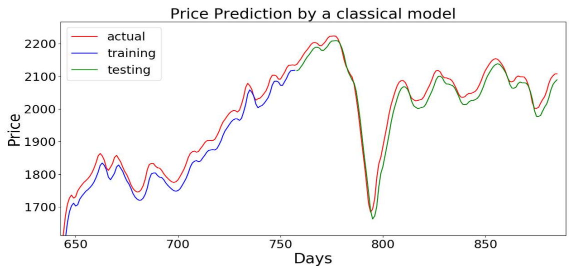

Figure 2: Result of a classical model. Its predicted values tend to be behind the actual ones in time, which is the problem of time lag.

3.1 Trend Price Prediction

In this part, our task is to make a rough estimation for the next day's recycling price, getting values around the actual ones. And when we get the predicted trend values, the residual values are obtained. There are many alternative methods and we choose MA model.

MA model is a simple way but can bring a similar effect in this task, compared with other complex models. It can get predicted values close to the actual ones but there is a distance between them, just like the predicted ones always behind the actual ones. It is the problem of time lag, which is common in the prediction using a single classical model.

Here is the basic idea of MA model. It makes prediction based on a known subsequence before the target time, which can eliminate the influence of random fluctuations and reflect the trend of time series changes. To be specific, in a certain time series, the average value of the data points in a fixed time length is taken as the predicted value of the next time point. The calculation of MA is as follows:

$$\hat{x}_{t+1} = \frac{S_{t-w+1,t}}{w} = \frac{x_{t-w+1} + x_{t-w+2} + \cdots + x_t}{w} \tag{2}$$

where $\hat{x}_{t+1}$ is the predicted value at the next time t+1, w is the length of the subsequence involved in the calculation of MA.

3.2 Residual Price Prediction

Residual price prediction is the focus of our model, aiming to eliminate the distance of the values predicted by MA model and the actual ones. So that our model can solve the problem of time lag and achieve a higher prediction accuracy.

In detail, we use EMD-LSTM model to predict the residual price sequence obtained by the first part. EMD method is used to decompose the residual price data and LSTM networks are used to make prediction for the component sequences.

3.2.1 EMD Method. EMD method is a signal decomposition method in time-frequency domain. It does not require pre-analysis of signals and can decompose a complex signal into a series of simpler and more stable components adaptively. It is considered to be suitable for analyzing nonlinear and non-stationary signals. In view that the residual price sequence reflects the short term volatility, which is unsteady and nonlinear, EMD method is adopted to process it first. It can help improve the performance of the LSTM network.

EMD method decomposes a complex time series into a finite number of Intrinsic Mode Functions (IMFs) and a residue. The IMFs are more stable and represent different levels of information in the original data. For time series $x(t)$, the steps of EMD are as follows[8]: (1) We calculate the upper and lower envelop of the initial sequence by getting all the local maximum and minimum values. (2) We calculate the average of the upper and lower envelop to obtain $m_1(t)$. (3) We get an intermediate signal $h_1(t)$ by using $x(t)$ minus $m_1(t)$. If $h_1(t)$ is not stationary, the above steps (1) - (3) will be repeated until it is stable. Then we can obtain the first IMF $c_1(t)$. (4) With $x(t)$ minus $c_1(t)$, a residual sequence $r_1(t)$ is obtained. Taking $r_1(t)$ as the initial sequence, the whole process is repeated until the last residual sequence $r_n(t)$ cannot be decomposed. Thus, the initial time series can be decomposed into:

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t) \tag{3}$$

3.2.2 LSTM Network. We use LSTM networks[9] to make prediction for those IMFs and residue. The LSTM network is a modified version of Recurrent Neural Network (RNN), overcoming some drawbacks of RNN. It has been proved effective for processing sequence data and can learn long-term dependency information. It is a good choice for the nonlinear residual price sequence. Thus, several LSTM networks with same structure are used to process the IMFs and residue separately to make prediction for each component.

The LSTM network has a special component called memory unit, which can determine whether some information is useful or not. A memory unit includes a memory cell and three gates: an input gate, a forget gate and an output gate. The input gate determines what new information should be stored in the cell state. There are two parts: tanh function produces a new candidate vector $\tilde{C}_t$ and sigmoid function gets a weight value i_t which decides the amount of information maintained. The formula is as follows:

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{4}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{5}$$

And then the forget gate and input gate is combined to update the memory cell state. The old state C_{t-1} is multiplied by f_t to decide the stored previous information. The new candidate vector $\tilde{C}_t$ is multiplied by i_t to decide the new information. The formula is as follow:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{6}$$

The output gate determines the final output value of the memory cell. Sigmoid function is used to decide the weight value o_t of the information filtering. And then o_t multiply with C_t which is processed by tanh function to obtain the final result h_t. The formula is as follows:

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \tag{7}$$

$$h_t = o_t * \tanh(C_t) \tag{8}$$

All the predicted component values obtained by the LSTM networks are added to get the final predicted residual values of the recycling price.

At last, we add the predicted trend values and residual values to get the results of the recycling price prediction.

4 Experiments

4.1 Dataset and Experimental Settings

Our dataset for this task comes from the website named 91 Regeneration[1]. It is the collection of daily recycling price data of four categories of scrap iron in Suzhou from 2015-08-03 to 2019-04-12. We will take one category as an example to illustrate in the following.

For our model, we use some means of data preprocess to improve the performance. In detail, we use the wavelet denoising method[10] to preprocess our data before the part of MA. And we normalize the data got from EMD and reconstruct them to meet the requirement of the LSTM network input. The input size of the LSTM networks is (N, L, 1), where N is the number of samples and L is the length of time steps used to make prediction. Moreover, we apply the dropout strategy and L2 regularization to mitigate overfitting of the network.

We implement the MA-EMD-LSTM with Keras and optimize it using Adam with a batch size of 32 and a learning rate of 0.005. For MA model, the parameter of moving window length w should not be too large. If the w is too large, the problem of time lag will be serious. And the time step length L of the LSTM network input also should not be too large. Because the data too long ago could bring extra noise to our model. So we search the optimal w and L within the ranges [1, 2, 3, 4, 5] and [2, 4, 6, 8, 10] respectively. Further, we search the optimal hyper-parameters of the LSTM network empirically. we tune the hidden neuron number n within [4, 8, 16, 32], the dropout rate p within [0.1, 0.2, 0.3, 0.4, 0.5], and the L2 regularization rate reg with [0.01, 0.05, 0.1]. The result is summarized in Table 1.

Table 1: Hyper-parameter settings

Hyper-parameter	Value
moving window length w	2
time step length L	8
hidden neuron number n	16
dropout rate p	0.2
L2 regularization rate reg	0.1

4.2 Performance Evaluation

In this part, we will introduce the evaluation indicators for model performance and some classical prediction models. we take their performance as the baseline of our model.

In order to evaluate the performance of different models, we select two common indicators, root mean squared error which is denoted as RMSE and coefficient of determination which is denoted as R^2. RMSE is usually used to measure the differences between the predicted values and the actual values. When RMSE is smaller, it means that the model's prediction accuracy is higher. R^2 provides a measure of how well a model can fit the actual values and make prediction. It is similar to the indicator accuracy in classification tasks. R^2 is between 0 and 1, where the larger R^2 means the better performance of the model. The formula of RMSE and R^2 are as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \tag{9}$$

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^{N}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{10}$$

where $\hat{y}_i$ is a predicted value, y_i is an actual value and $\bar{y}$ is the mean of data.

In addition, in order to evaluate the effect on the renewable resource investment in reality, we simulate the process of the investment to calculate earnings in a period. We take a ratio between the earnings guided by the predicted price and those guided by the actual price as an indicator of evaluation, denoted as γ. The larger the indicator γ, the better the performance of a prediction model.

We compare the performance of our proposed model with some baselines for evaluation. The methods to be compared are the SVR model and the single LSTM network. After repeated testing, the hyper-parameter setting of SVR is C=1e3, gamma=0.5. The LSTM network's setting is same as those LSTM networks in our model.

4.3 Results and Analysis

In this part, we will give the results of our experiment and make some analysis. First, the results of the baselines and our model are shown in figure 3, 4, 5.

It can be seen from these figures that our model has a better performance than the other two classical models. To be specific, the values predicted by SVR or a single LSTM network tend to be behind the actual values, which is the problem of time lag. Since our task is to predict the recycling price one day ahead, this problem of baselines will bring poor performance in practical applications. However, our model does not have this problem so that it can get a higher prediction accuracy.

[1] http://www.zz91.com

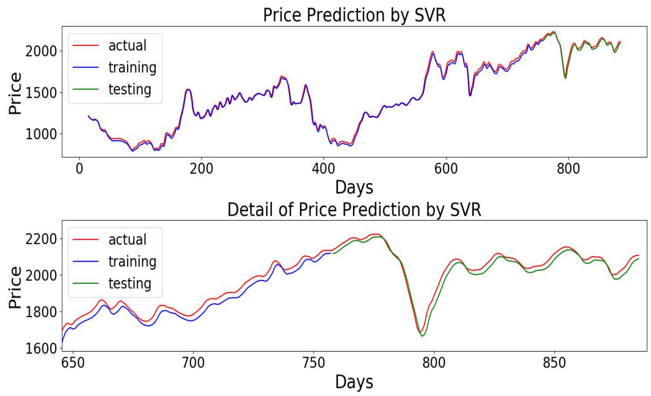

Figure 3: Result of SVR model. It has the problem of time lag obviously both on training and testing data

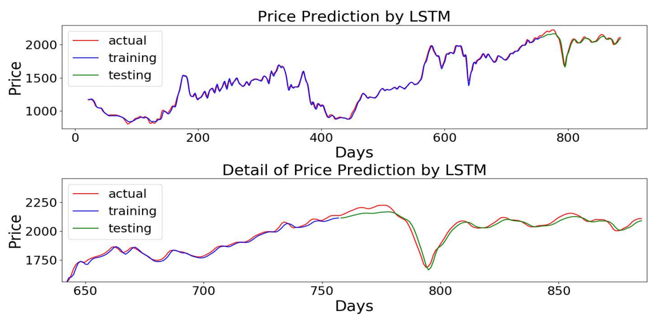

Figure 4: Result of single LSTM network. It also has the problem of time lag and has a big error on testing data.

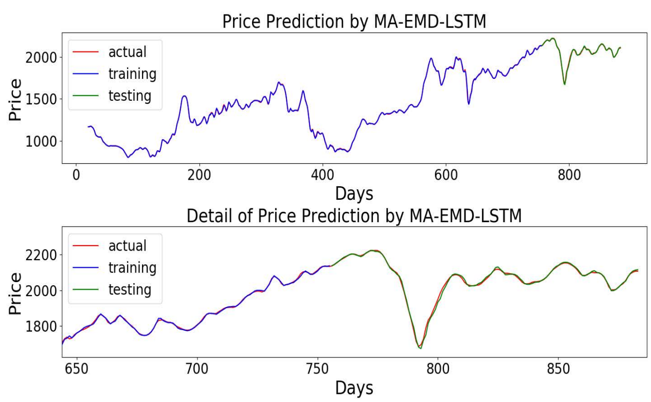

Figure 5: Result of our MA-EMD-LSTM model. Unlike the two classical models, it works very well on both training and testing data, without the problem of time lag.

Further, we draw a comparison between the baselines and our model in the three evaluation indicators, shown in Table 2.

Table 2: Comparison in evaluation indicators

Model	RMSE	R^2	γ
SVR	26.4189	0.9349	0.0375
LSTM	31.6623	0.9152	0.0418
MA-EMD-LSTM	**7.3470**	**0.9954**	**0.8039**

From the table, we can see clearly our model get the best score three indicators. It makes the value of RMSE below 10, 72.19% lower than SVR model and 76.80% lower than single LSTM network. The indicator R^2 of MA-EMD-LSTM model is very close to 1, which directly shows the highest accuracy and best performance of our model. Further, the indicator γ is around 0.8, much larger than that of SVR or LSTM model which is only below 0.1. It means our model also has a good effect in the actual application.

5 Conclusion

In this paper, we introduce an approach for recycling price prediction of renewable resources. We propose a combined model based on the method of MA and EMD and the LSTM network to address the problems in this task. Our model divides the task into trend prediction and residual prediction so that the problem of time lag can be solved. We use the method of MA to predict trend values. And the residual prediction is the key point of our model. EMD is used to preprocess the input data of LSTM networks in order to improve the performance of the neural network. We carried on some experiments to evaluate our model. The experimental results validate the effectiveness of our approach.

In future, we plan to explore the following directions: 1) we are interested in the contribution of MA and EMD to our model respectively, which could help to improve our model. 2) we plan to apply our model to a more diverse dataset to further prove the effectiveness of our approach. 3) we will conduct the qualitative analysis on our experimental results.

REFERENCES

[1] Shaufique F. Sidique, Satish V. Joshi, Frank Lupi. (2010). Factors influencing the rate of recycling: An analysis of Minnesota counties. Resources, Conservation and Recycling, 54(4): 242-249.

[2] Doron Lavee, Uri Regev, Amos Zemel. 2009. The effect of recycling price uncertainty on municipal waste management choices. Journal of Environmental Management, 90(11): 3599-3606.

[3] Gordon Werner, Shanchieh Yang, and Katie McConky. 2017. Time series forecasting of cyber attack intensity. In Proceedings of the 12th Annual Conference on Cyber and Information Security Research (CISRC '17). ACM, New York, NY, USA, Article 18, 3 pages. DOI: https://doi.org/10.1145/3064814.3064831

[4] Xianjing Wang, Jonathan Liono, Will McIntosh, and Flora D. Salim. 2017. Predicting the city foot traffic with pedestrian sensor data. In Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous 2017). ACM, New York, NY, USA, 1-10. DOI: https://doi.org/10.1145/3144457.3152355

[5] Hsueh-Yi Lu and Lu-Ting Kuo. 2018. Development of a Smart Drainage Control System Using Support Vector Regression. In Proceedings of the 2nd International Conference on Medical and Health Informatics (ICMHI '18). ACM, New York, NY, USA, 212-217. DOI: https://doi.org/10.1145/3239438.3239441

[6] Michael Schultz and Stefan Reitmann. 2018. Prediction of aircraft boarding time using LSTM network. In Proceedings of the 2018 Winter Simulation Conference (WSC '18). IEEE Press, Piscataway, NJ, USA, 2330-2341.

[7] Weixi Gu, Zimu Zhou, Yuxun Zhou, et al. 2017. Predicting Blood Glucose Dynamics with Multi-time-series Deep Learning. In Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems (SenSys '17), Rasit Eskicioglu (Ed.). ACM, New York, NY, USA, Article 55, 2 pages. DOI: https://doi.org/10.1145/3131672.3136965

[8] Norden E. Huang, Zheng Shen, Steven R. Long, et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings Mathematical Physical & Engineering Sciences, 454(1971):903-995

[9] Mohammad Assaad, Romuald Bone´ and Hubert Cardot. (2008). A new boosting algorithm for improved time-series forecasting with recurrent neural networks. Information Fusion, 9(1):41-55

[10] Rami Cohen. (2012). Signal denoising using wavelets. Project Report, Department of Electrical Engineering Technion, Israel Institute of