Identifying robust landmarks in feature-based maps

Julie Stephany Berrio 1, James Ward 1, Stewart Worrall 1, Eduardo Nebot 1

Abstract—To operate in an urban environment, an automated vehicle must be capable of accurately estimating its position within a global map reference frame. This is necessary for optimal path planning and safe navigation. To accomplish this over an extended period of time, the global map requires long term maintenance. This includes the addition of newly observable features and the removal of transient features belonging to dynamic objects. The latter is especially important for the long-term use of the map as matching against a map with features that no longer exist can result in incorrect data associations, and consequently erroneous localisation. This paper addresses the problem of removing features from the map that correspond to objects that are no longer observable/present in the environment. This is achieved by assigning a single score which depends on the geometric distribution and characteristics when the features are re-detected (or not) on different occasions. Our approach not only eliminates ephemeral features, but also can be used as a reduction algorithm for highly dense maps. We tested our approach using half a year of weekly drives over the same 500 metre section of road in an urban environment. The results presented demonstrate the validity of the long term approach to map maintenance.

I. INTRODUCTION

The accurate estimation of vehicle pose within a global environment is a fundamental requirement for the navigation of an autonomous vehicle. Urban vehicles must be able to operate in areas where GNSS cannot provide accurate localisation, or at times does not provide any position information at all [1]. This means that an alternative method of localisation is required, and a popular approach is to perform pose estimation based on the mapping of distinctive landmarks in the operational area.

The generation of such maps is challenging as it requires the initial acquisition of sensor data, information analysis, and continuous maintenance [2]. Feature-based maps provide a compact representation of the environment. The main assumption in their construction is that the world is static, and hence the features do not change over time. Also, when the map is first generated every feature observed by the perception system is incorporated as a map landmark, whether or not it represents a static or dynamic object. Consequently, maps should be updated as the environment changes to aid not only localisation, but motion planning in dynamic environments [3].

The initial map features and their properties (dynamic/static, feature class) obtained from a SLAM type algorithm depend on the configuration parameters that control the sensitivity and quality of the feature detector. For example, if the detector is restricted to selecting only features that are located above 3 meters from the ground plane, it is far more likely that the detected features correspond to static objects. The problem with using a very restrictive detector is that there is a corresponding reduction in the number of features. With fewer features to enable the localisation, there is an increased chance of poor map matching and failures in the data association procedure.

The maintenance process generally includes long term data collection from a vehicle travelling within the same area on different days, times of day or seasons. The post processing of information should identify and incorporate any environmental changes into the map representation. In a typical setup, a local map is created and stored for the current session while the vehicle is navigating [4]–[6]. In the case that a known place is visited again, new observations are used to improve the map by adding new features [7] and incrementally updating the feature map.

For the operation of a SLAM type approach, incorporating more features into the state vector can lead to an unmanageable amount of information. This is compounded when using a map with features belonging to objects that are no longer in the scene. In some approaches, the map is stored and queried in order to retrieve only the features which could be associated with the current sensor information. When
a system is bandwidth limited the amount of data to be transferred and computed has to be taken into account. In both cases is important determine when the map features no longer correspond to objects in the environment, not only to keep the compactness and reliability of the map, but also to simplify the data association problem.

In this paper we present a method to update the information contained in a feature map using a process that estimates a feature score. This approach can also be used as a method to downsample the map. A regression model is trained with a data set labelled using empirical probability estimation. The data set consists of a group of identified predictors based on the information collected over 6 months using an electric vehicle. This vehicle is equipped with various sensors, and was driven to follow approximately the same trajectory once per week within a predefined area. A prior feature map was created from a dataset collected during a single drive around this area.

Our approach achieved both the scoring and selection of reliable features and also the downsampling of the map while maintaining a guaranteed coverage of important features in areas with scarce landmarks. We tested our approach using a map generated using detected pole and corner features as inputs to an Extended Kalman filter (EKF) for simultaneous localisation and mapping [8]. The map was used to localise during subsequent drives, and the testing was carried out using the new data. Our evaluation method consists of localising the vehicle within the resulting map and calculating the performance of the localisation algorithm by computing the magnitude of the maximum vehicle state covariance. In the same way, the map downsampling was assessed by the selective dropping of landmarks (based on our score) from the original map.

In the next section, we introduce a review of the related work for map updating through the removal of features that no longer exist. In Section III, we describe the components of our approach and the methodology used to build the model. The experiments, evaluation and outcomes are presented in Section IV.

II. RELATED WORK

Changes in the perceived environment can be caused by environmental conditions (weather, trees foliage due to the seasons), time of day (shadows, illumination) and structural variations (new/demolished buildings, dynamic objects taken into account when the map was made). Some of these changes can be managed by just adding new information into the map, and relying on the data association algorithm to ignore outliers produced by these changes.

Different works have addressed the problem of removing dynamic or non-existing features from feature maps in both vision-based and lidar-based systems. For visual localisation, to estimate the robot pose within an existing map it is necessary to retrieve the most reliable features (due to and bandwidth limitations) in the local map from the map storage to calculate the current pose [9]. A diverse range of approaches have been used in order to reduce the size of a feature map. In [10], the authors presented an approach to choosing landmarks that are considered valuable for localisation based on statistical measures. In this work, each landmark is scored as a function of the times it has been observed. In [11], an algorithm that alternates between on-line place recognition and off-line map maintenance is presented and evaluated with the aim of producing a reliable fixed size map. The same authors in [7] propose a method for reducing the amount of map data that needs to be transferred by making use of predictors that include the distance travelled while observing a landmark, mean re-projection error and the classification of the descriptor appearance. Stability of visual features has also been explored, e.g. by assessing their distinctiveness and the robustness of image descriptors [12], or uniqueness of the descriptors [13].

A different approach is presented in [14] where the dynamic environment is represented by occupancy grids. Each cell has an associated state transition probability. Another strategy for removing old objects is based on associating a persistence time measure to each feature. This approach assigns a set life time to each feature each time it is re-observed [15]. If by the end of this time span the feature has not been matched with any observation, it is removed from the map [16]. A recent work described as frequency map enhancement [17], [18] develops a spatio-temporal representation to describe the persistence and periodicity of the individual cells/features allowing the future calculation of an occupancy/occurrence prediction probability.

In contrast to the visual based mapping approaches, our model is generic and can be used to map any kind of feature/landmark in 2D or 3D. This is because it is based on predictors that are not attached to specific feature properties, instead based only on the geometry as measured from the dataset. This methodology was validated with real data using a prior 2D map.

III. METHODOLOGY

The measure of importance given to a landmark for localisation purposes is established based on a combination of different variables. The objective of this paper is to derive,
examine, analyse and evaluate these importance variables. The variables are related to the observability and distribution of the landmarks, leading to a relevance score.

Different predictors are proposed, selected, adjusted and later combined to evaluate the capacity of each feature to be used for localisation in future datasets. We defined some notations as follows:

- \([x_{fn}, y_{fn}]\) denotes the location of the landmark \(n\) in the map global frame.
- \([x_v, y_v]\) denotes the location of the vehicle in the map global frame.
- \(\theta_{fn}\) is the detection angle of the landmark \(n\) in its own local frame when the vehicle is located in \([x_v, y_v]\).
- \(r_{\theta fn}\) is the distance between the vehicle and the landmark when the last time it was detected at the angle \(\theta_{fn}\).
- \(\theta_{s_{n_{max}} fn}\) denotes the spanned angle of the landmark detection based on the vehicle’s trajectory.
- \(df_{n-m}\) denotes the distance between the landmarks \(n\) and \(m\) respectively.

### A. Predictor variables

1) **Number of detections**: By considering variables associated with the feature’s influence on the quality of localisation and its persistence during the vehicle’s operation, it is clear that the number of times \(n_{fn}\) the landmark \(n\) has been observed is directly related to the usefulness of the feature. This predictor has been used extensively in map downsampling approaches [10], [19]. The value as a predictor is based on the assumption that the more matches of the landmark \(n\), the greater the expectation that it will be detected again and consequently be used by the localisation algorithm. Any time a landmark is matched to the prior map, the corresponding feature counter \(n_{fn}\) is increased by one, regardless of the vehicle’s position or speed.

2) **Maximum detected spanning angle**: This variable corresponds to the maximum angle covered by all detections/matches achieved from the moving vehicle to each landmark. The fact the landmark can be seen from a comparably extended set of angles illustrates the relevance of that landmark to localize the vehicle in the surrounding area. To every landmark, a vector of discretized angles \(\theta_{df_{n}}\) is assigned, in our case with 1° resolution, at the moment the landmark is successfully matched, the angle to the vehicle \(\theta_{fn}\) is calculated and rounded, and the corresponding vector position is marked indicating that a successful feature match was achieved from that angle. Thus, the maximum detected spanned angle \(\theta_{d_{m_{fn}}}\) of the landmark \(n\) is calculated as the sum of all components within \(\theta_{df_{n}}\).

\[
\theta_{d_{m_{fn}}} = \text{sum}(\theta_{df_{n}})
\]  

3) **Maximum length driven while observing the landmark**: A predictor derived from \(\theta_{d_{m_{fn}}}\) is the maximum length driven while the landmark is being observed \(Ml_{fn}\). The length of the path from where the landmark can be observed is taken into consideration. The \(\theta_{d_{m_{fn}}}\) also depends on the environment structure (open area or with the presence of several obstacles) and the vehicle’s trajectory (rectilinear or curvilinear) within the map. Similarly to the calculation of \(\theta_{d_{m_{max}}{fn}}\), a vector of ranges associated to each discrete angle \(mR_{fn}\) is designated to each landmark. When the landmark is re-observed, the distance from the vehicle to the landmark \(r_{\theta fn}\) is computed and compared with any previous value belonging to the same detection angle \(\theta_{fn}\). The maximum value is then registered in the vector position corresponding to \(\theta_{fn}\). The \(Ml_{fn}\) is determined as the length of the arc described by the vehicle. This is calculated as the dot product between the vector \(mR_{fn}\) and the vector of discrete unit angles \(A_{r}\) for which all components are equal to \(pi/180\) (1° angle resolution).

\[
Ml_{fn} = mR_{fn} \cdot A_{r}
\]  

4) **Maximum area of detection**: In a similar manner, the Maximum area of detection \(Ma_{fn}\) was calculated. In this predictor we compute the total detection area of the landmark by assuming that it could be detected by any of the geographical points between it’s position \((x_{fn}, y_{fn})\) and the vehicle’s pose \((x_v, y_v)\). \(Ma_{fn}\) is estimated as half the dot product between the pairwise power of the vector \(mR_{fn}\) and \(A_{r}\).

\[
Ml_{fn} = \frac{1}{2} mR_{fn}^2 \cdot A_{r}
\]
5) Maximum possible spanned angle: Per every feature \( n \) and by looking for the vehicle’s vector poses within a radius of 30 meters on the feature frame, we calculated the maximum possible spanned angle \( \theta_{s_{\text{max}},n} \). This predictor depicts the landmark’s potentiality for localisation. It is not susceptible to environmental changes that may occlude the landmark since it is assumed that the vehicle navigates in an open space, hence \( \theta_{s_{\text{max}},n} \) depends only on the vehicle’s trajectory.

6) Concentration ratio: The spatial distribution of landmarks within the map should be considered in any feature map generation algorithm. Areas with few landmarks for localisation should keep as many of the features as possible, whereas areas with dense features can be reduced. To measure the weight of each landmark within its surroundings, we introduce a predictor called concentration ratio \( Cr_n \). The concentration ratio for a given feature indicates the density of surrounding features used for localisation. To calculate the concentration ratio \( Cr_n \) for every landmark \( n \) in the map, all distances \( df_{n-i} \) to the landmarks \( i \) located within a 30 metre radius of \( n \) are calculated. The value of \( Cr_n \) is then estimated as the division between the furthest landmark \( \max(df_{n-i}) \) from \( n \) and the sum of the distances \( df_{n-i} \).

\[
Cr_n = \frac{\max(df_n)}{\sum df_{n-i}} \tag{4}
\]

A low value of \( Cr_n \) indicates a higher density of features for localisation compared with a concentration ratio approaching 1 which indicates the landmark density is sparse.

B. Scoring function

The scoring function aims to estimate the relevance to localisation for each landmark across a number of datasets as a function of each of the predictors. We propose to use a regression algorithm to adjust the coefficients of the identified predictors calculated using a training dataset classified by its empirical probability. The empirical probability function can be defined as the vector of frequencies of each landmark detection when considering the dataset [20]. This vector is normalized using the percentage of the datasets in which the landmark has been re-observed. Each dataset was registered in the same area used to create the initial feature map. As the vehicle was driven approximately along the same trajectory each time, persistent landmarks from the map should be recognized. The localisation algorithm was modified to retrieve the details of the matched landmarks each time a detection was performed.

The predictors need to be standardised before they can be incorporated into the scoring function. This is required as the scale of the predictors varies and for the regression model we are proposing, normalisation of the scale is needed for each parameter to contribute equally. After this pre-processing step, we perform use a ridge and lasso model to minimise the regression coefficients. This is based on a sum of the squared coefficients and the sum of the absolute coefficients. The predictors with lesser significance to the result have the corresponding coefficients set to near zero. It is critical for the analysis of the predictors that they can be selected, removed or transformed to formulate a reliable model.

After the predictor coefficients were selected, we use a cross-validated elastic net regularized regression method [21] to fit our model. The method is switched between Lasso (L1) and Ridge (L2) regularization methods to overcome the limitations of each approach based on an alpha parameter. Elastic net is suitable for this case where some predictors are strongly correlated.

IV. RESULTS

We tested our approach using a prior feature map generated from a single dataset Fig. 4. An electric vehicle (EV) equipped with a 16 beam Velodyne laser sensor was driven in a clockwise direction in front of the main quadrangle of the University of Sydney to collect the dataset used to build the initial feature-map. No special conditions were enforced for the area of the collected dataset, meaning vehicles, pedestrians and cyclists were also observed by the laser. The length of the track was around 500 metres.

The features extracted from the point cloud corresponds to two feature categories, poles and corners, which are presumed to belong mainly to static objects. However, corners of parked and moving vehicles (especially trucks) and pedestrians were sometimes detected as poles and included as observations. An extended EKF-SLAM algorithm was used to build the prior feature map shown in Fig. 1.

Fig. 4: A 500m track length dataset was captured once per week over a six month period at different times of day.

Over the course of six months, data collection was performed around the University of Sydney campus on a weekly basis on different days of the week and hours of the day. During this period different construction works were carried out in the mapped area, causing the disappearance of objects in the zone. This information obtained over this half year time period was used to build and test our model.

The localisation algorithm using the prior map estimated the global pose of the vehicle based on an iterative closest point (ICP) data association algorithm, which identified the pose corresponding to the current observations. An offline algorithm was implemented to compute each of the predictors based on the localisation algorithm output (vehicle pose and observed features).
For predictor selection we used ridge and lasso regularisation models, and checked which coefficients were set to zero. For the lasso case, the coefficient belonging to the number of detections of each landmark was set to zero due to its high correlation with other predictors, having a maximum correlation coefficient $R^2$ of 0.96 with the detected spanned angle. For the ridge model, the coefficients set to zero corresponded to Maximum possible spanned angle and the concentration ratio. Since we want our model to include the concentration ratio so that areas with a lower density of landmarks will be maintained, we used a non-linear transformation formed by the multiplication of the coefficient ratio and the number of detections.

Table I shows the $R^2$ between the predictors and the empirical probability used to label the data.

| Predictor            | $R^2$   | Predictor           | $R^2$   |
|----------------------|---------|---------------------|---------|
| Number of views      | 0.2352  | Maximum Angle       | 0.0599  |
| Spanned Angle        | 0.2947  | CR                  | 0.0184  |
| Track Length         | 0.2805  | CR*N Views          | 0.2523  |
| Total Area           | 0.2562  |                     |         |

TABLE I: Predictors’ coefficients of correlation with the labelled output.

Having selected and transformed the predictors, we used a cross-validated elastic net regularized regression method to tune the $\alpha$ parameter and select the coefficients corresponding to $\lambda$ with minimum cross-validation error plus one standard deviation. The achieved coefficient of correlation $R^2 = 0.76$, which implies that three quarters of the variance is related to the predictors.

![Fig. 5: Trace plot of coefficients fit by elastic net. The blue dotted line denotes the $\lambda$ with minimum cross-validation error plus one standard deviation.](image)

A map with the previous selected landmarks, Fig.7, has 35% less features than the initial map. Through a visual inspection of the zone and the map itself, we corroborated the correspondence to the ground truth. Our approach was able to identify and remove features in the middle of the road, pedestrian paths, parking areas, and other features difficult to identify by the sensors. Also, landmarks that were very high from the ground and inside of buildings were likely to be pruned.

![Fig. 6: Predictions histogram and distribution.](image)

![Fig. 7: Map built with identified static and semi-static landmarks in green, discarded landmarks in red.](image)

Since the regression predictions score the landmarks for localisation, we tested our downsampling capability and compare it with the approach in [11] which only takes into account the travelled distance while observing the landmark. The dataset used for the experimental results come from the later weeks in the half year of weekly datasets, and therefore was not included in the initial model fitting. Fig.8 show the outcome of this comparison. We started by dropping the least valuable 20% of the features. The maximum covariance magnitude remains the roughly same for both methods. After 20% drop rate this value for our proposed method is always below the “distance travelled while observing” metric, and is almost constant until a drop rate of 70%. It is important to note that the evaluation of the final map in Fig.7 corresponds to 35% drop rate, which is where the maximum covariance magnitude shows a negligible increment in contrast to the entire map.

Both methods will lead to failures of the localisation algorithm once the feature map is very sparse. This failure
occurs when the drop percentage is 75% for the method based on the track length and 85% for the proposed method. This demonstrates that our method is better at predicting the importance of features for localisation and can correctly prioritise the removal of less important features.

V. CONCLUSION AND FUTURE WORK

In this paper we presented an approach to evaluate landmarks based on the likelihood of being persistently re-observed over future visits to the mapped area. Our approach allows us not only to discard unstable landmarks and have a map more suited to perform localisation, but also score each landmark so that the map can be reduced while maintaining reliable localisation.

The scoring method is based on an elastic net regression model which incorporates a range of predictor variables that are related to the capability of each landmark to improve the localisation of the vehicle. The feature concentration ratio was included as a novel predictor to allow the system to measure the degree of feature sparseness related to localisation in the neighbouring area and consequently give it a larger weight in the scoring function. Even though maximum possible spanned angle predictor was not used in this model, we believe it could be a valuable predictor to include in cases where the vehicle/robot can travel the mapped area in two different directions. This is because the feature detector does not detect features uniformly within the sensor’s coverage area, and that the environment was crowded with pedestrians and vehicles. In addition, there were various sources of occlusions that would commonly occur. This can result in the vehicle travelling within the map but not estimating the same track that was travelled when the original map was built. For this paper, the data collection was done in a single direction and this predictor was discarded. In future work we intend to include this metric in order to assess the contribution when the vehicle is operated in the cases mentioned.

We demonstrate the downsampling ability of the model by comparing its performance with a state of the art model. Our model achieves better results for localisation purposes at all downsampling ratios.

ACKNOWLEDGMENT

This work has been funded by the ACFR, the University of Sydney through the Dean of Engineering and Information Technologies PhD Scholarship (South America) and the Australian Research Council Discovery Grant DP160104081 and University of Michigan / Ford Motors Company Contract “Next generation Vehicles”.

REFERENCES

[1] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt, M. Sokolsky, G. Stunek, D. Stavens, A. Teichman, M. Werling, and S. Thrun, “Towards fully autonomous driving: Systems and algorithms,” in 2011 IEEE Intelligent Vehicles Symposium (IV), June 2011, pp. 163–168.
[2] S. Krutti, S. Fallah, K. Katsaros, M. Ducati, F. Mccullough, and A. Mouzaki, “A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications,” IEEE Internet of Things Journal, vol. 5, no. 2, pp. 829–846, April 2018.
[3] F. Pomerleau, P. Krši, P. Colas, P. Furgale, and R. Siegwart, “Long-term 3d map maintenance in dynamic environments,” in 2014 IEEE Intelligent Conference on Robotics and Automation (ICRA), May 2014, pp. 3712–3719.
[4] R. Guo, F. Sun, and J. Yuan, “Icp based on polar point matching with application to graph slam,” in 2009 International Conference on Mechatronics and Automation, Aug 2009, pp. 1122–1127.
[5] D.-I. Kim, H. Chae, J. B. Song, and J. Min, “Point feature-based outdoor slam for rural environments with geometric analysis,” in 2015 12th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Oct 2015, pp. 218–223.
[6] P. Sirianawan, M. D. Prasanjith, and D. Wang, “3d feature points detection on sparse and non-uniform pointcloud for slam,” in 2017 18th International Conference on Advanced Robotics (ICAR), July 2017, pp. 112–117.
[7] M. Dymczyk, T. Schneider, I. Glütschenksi, R. Siegwart, and E. Stumm, “Erasing bad memories: Agent-side summarization for long-term mapping,” in 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Oct 2016, pp. 4572–4579.
[8] M. W. M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba, “A solution to the simultaneous localization and map building (slam) problem,” IEEE Transactions on Robotics and Automation, vol. 17, no. 3, pp. 229–241, June 2001.
[9] S. Lynen, T. Sattler, M. Bosse, J. A. Hesch, M. Pollefeys, and R. Siegwart, “Get out of my lab: Large-scale, real-time visual inertial localization,” in Robotics: Science and Systems, 2015.
[10] P. Mühlfellner, M. Bürki, M. Bosse, W. Derendarz, R. Philipsen, and P. Furgale, “Summary maps for lifelong visual localization,” Journal of Field Robotics, vol. 33, no. 5, pp. 561–590, 2016, this work is supported in part by the European Community’s Seventh Framework Programme (FP7/2007-2013) under Grants No. 269916 (V-Charge) and No. 610603 (EUROPA2).
[11] M. Dymczyk, S. Lynen, T. Cieslewski, M. Bosse, R. Siegwart, and P. Furgale, “The gist of maps - summarizing experience for lifelong localization,” in 2015 IEEE International Conference on Robotics and Automation (ICRA), May 2015, pp. 2767–2773.
[12] G. Carneiro and A. D. Jepson, “The quantitative characterization of the distinctiveness and robustness of local image descriptors,” Image and Vision Computing, vol. 27, no. 8, pp. 1143 – 1156, 2009.
[13] Y. Verde, K. M. Xi, P. Fua, and V. Lepetit, “Tilde: A temporally invariant learned detector.” IEEE Computer Society, 2015, pp. 5279–5288.
[14] D. Meyer-Delius, M. Beinhofer, and W. Burgard, “Occupancy grid models for robot mapping in changing environments,” in Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence, ser. AAAI’12. AAAI Press, 2012, pp. 2024–2030. [Online]. Available: http://dl.acm.org/citation.cfm?id=2900929.2901014.
[15] D. M. Rosen, J. Mason, and J. J. Leonard, “Towards lifelong feature-based mapping in semi-static environments,” in 2016 IEEE International Conference on Robotics and Automation (ICRA), May 2016, pp. 1063–1070.
[16] F. Nobre, C. Heckman, P. Oszog, R. Wolcott, and J. Walls, “Online probabilistic change detection in feature-based maps,” in 2017 IEEE International Conference on Robotics and Automation (ICRA), May 2017, pp. 3661–3668.
[17] T. Krajnk, J. P. Fentanes, M. Hanheide, and T. Duckett, “Persistent localization and life-long mapping in changing environments using the frequency map enhancement,” in 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Oct 2016, pp. 4558–4563.

[18] T. Krajnk, J. P. Fentanes, J. M. Santos, and T. Duckett, “Fremen: Frequency map enhancement for long-term mobile robot autonomy in changing environments,” IEEE Transactions on Robotics, vol. 33, no. 4, pp. 964–977, Aug 2017.

[19] J. W. Kaeli, J. J. Leonard, and H. Singh, “Visual summaries for low-bandwidth semantic mapping with autonomous underwater vehicles,” in 2014 IEEE/OES Autonomous Underwater Vehicles (AUV), Oct 2014, pp. 1–7.

[20] M. Vani and K. Urbaniec, “Employing bayesian networks and conditional probability functions for determining dependences in road traffic accidents data,” in 2017 Smart City Symposium Prague (SCSP), May 2017, pp. 1–5.

[21] H. Zou and T. Hastie, “Regularization and variable selection via the elastic net,” Journal of the Royal Statistical Society, Series B, vol. 67, pp. 301–320, 2005.