Roof-Crush Protection Design of Automotive Bodies Using Clustering and Pattern Recognition

Yong-Sang Shin, Hyo-Jun Eun, Yong-Ju Chu and Seung-Yop Lee *

Department of Mechanical Engineering, Sogang University, Seoul 04107, Korea; wmsk53@gmail.com (Y.-S.S.); ehj93@naver.com (H.-J.E.); yjchu@katech.re.kr (Y.-J.C.)
* Correspondence: sylee@sogang.ac.kr; Tel.: +82-2-705-8638

Received: 4 March 2019; Accepted: 2 April 2019; Published: 5 April 2019

Abstract: Computer-aided engineering (CAE) tools play an indispensable role in the vehicle development process. However, it is difficult to accurately predict the relationships and behavior of automotive bodies in vehicle crashes owing to high-order nonlinearity and numerous design variables of the automotive body structure. In this study, clustering and pattern recognition techniques were used to develop a novel optimization design of an automotive body considering roof crushing by vehicle rollover. The large-scale data were clustered to find the strong and weak clusters, and new response surface models were acquired by clustering analysis to achieve better performance than the response surface model of traditional optimization. For an efficient robust design, clusters with weak performance were excluded from the optimum solution. Finally, it was confirmed that the solutions by the proposed optimization technique were better than those obtained by the traditional optimum method based on a comparative analysis by various cluster combinations.

Keywords: clustering; pattern recognition; machine learning; optimum design; vehicle; roof crush

1. Introduction

International environmental regulations and fuel efficiency demands in developed countries are continually putting pressure on the automotive industry. For fuel efficiency improvement, weight-reduction technologies are relatively easy to apply and have the advantage of being applicable to both existing internal combustion engine vehicles and futuristic vehicles. Weight is a core factor that determines vehicle fuel efficiency; weight reduction is also associated with various other design issues, including structural rigidity, aerodynamic design, braking system, CO₂ reduction [1–3]. Because of this, the automotive industry has consistently applied weight reduction to various vehicle parts. As shown in Table 1, the body-in-white (BIW) accounts for most of the weight in a vehicle along with the power train and chassis, and it has a high potential for weight reduction [1,4,5]. A vehicle’s BIW includes parts that require excellent durability because they must be able to protect occupants by preventing deformation from being transferred to occupant space during accidents. Among vehicle collision accidents, damage caused by head-on collisions has typically led to the most severe injuries to occupants; however, such collisions have shown a decreasing trend over time by various safety technologies. However, loss of human life owing to vehicle rollover has shown an increasing trend. Furthermore, during head-on collisions, the bumper and engine parts are deformed themselves, and the collision energy is absorbed to reduce the shock and injury experienced by the occupants. However, in the case of a rollover, external impact is transferred directly to the occupant space, which leads to fatal injuries. In particular, rollover crashes are serious safety problems for light-weight vehicles [6]. For this reason, many studies have been conducted on the optimization of roof strength in regard to a simulated vehicle rollover [7–9]. Accidents, including vehicle rollovers, have high nonlinearity and it is difficult to predict the deformation behavior of parts over time after an accident. Presently,
when considering completely finished vehicles, competitive products that are durable and lightweight must be produced quickly, and this requires new optimization methods that use computer-aided engineering (CAE); this is different than what has been carried out in the past.

| System       | Major Components in System                                                                 | Mass Breakdown (%) |
|--------------|-------------------------------------------------------------------------------------------|-------------------|
| Body-in-white| passenger compartment frame, cross and side beams, front-end structure, roof structure, panels | 23–28             |
| Powertrain   | engine, transmission, exhaust system, fuel tank                                            | 24–28             |
| Chassis      | chassis, suspension, tires, wheels, steering, brakes                                       | 22–27             |
| Interior     | seats, insulation, trim, airbags                                                          | 10–15             |
| Closures     | front and rear doors, hood, lift gate                                                      | 7–8               |
| Miscellaneous| lightning, windows, glazing                                                               | 7–8               |

Data classification and clustering, pattern recognition, and machine learning are important problems in a variety of engineering and scientific disciplines, such as biology, psychology, medicine, marketing, computer vision, artificial intelligence, and remote sensing [10–14]. Numerous artificial intelligence techniques have been used to optimize various problems in the fields of mechanical and automobile engineering including sensing and control of unmanned vehicles [15,16], fault diagnosis of machineries [17,18], and manufacturing applications [19,20].

This study proposed a new optimization method that addressed the limitations of existing traditional methods. We created models according to design variable changes using methods of acquiring and optimizing regression models that employ these models, but combine machine-learning concepts such as clustering and pattern recognition in addition to traditional methods. This new method was compared to traditional optimization methods to prove its quality.

2. Modeling and Finite Element Analysis

2.1. Simplified FE Modeling

In this study, a simplified finite element (FE) model was created to optimize the BIW design process as shown in Figure 1, based on the open computer aided design (CAD) files of a 2012 Toyota Camry model [21]. Then, basic data were created by changing the design variables to apply machine-learning techniques such as clustering and pattern recognition with optimization [9]. The simplified BIW model consisted of the design variables shown in Table 2. To reduce the time spent on finite element analysis (FEA), all BIW elements were configured as beam elements. The roof panel was configured as a rigid body and connected to each part of the BIW. To simulate the roof impact caused by a rollover, a rigid body wall at an incline of 23° was made to collide obliquely with the upper part of the car at 700 mm/s, as shown in Figure 1. A boundary condition was configured in which the end of each pillar was completely fixed while the rigid body wall collided with the BIW [7,9]. In addition, SAE-AISI 1045 carbon steel (S45C) material was used for each part that makes up the BIW, and ABAQUS commercial software was used for the FEA.
Traditional optimization undergoes a process in which a regression model such as a response surface model (RSM) is set up by observing the responses according to design variables; and optimization techniques are used to predict the optimal values [25]. However, this study proposes a new optimization technique using machine-learning concepts including clustering and pattern recognition, in addition to the traditional optimization technique, as shown in Figure 2. Similar to the traditional technique, the maximum reaction force to the rigid barrier impact and the mass of the BIW model were set as responses, and the changes in responses in 2000 models were observed according to changes in the ten design variables listed in Table 2. After this, a concept known as displacement data was proposed in order to perform clustering based on the similarity of shape changes over time in the 2000 models. Displacement data implied a deformation displacement of nodes created by the BIW model when the rigid barrier pushed the BIW model and it moved. The data were created by extracting the displacement of all 203 nodes of the BIW model at six points when the rigid barrier moved 7.5, 22.5, 37.5, 52.5, 67.5, and 82.5 mm in the movement direction (y and z directions) [6]. Therefore, there were a total of 2436 (203 nodes × 2 directions × 6 timing) displacement datapoints for each analysis model. The resulting data structure is shown in Table 3.

### Table 2. Design variables of optimization for 2000 body-in-white (BIW) models.

| Parts                        | Variation Range |
|------------------------------|-----------------|
| Front/Middle/Rear pillar     | 10–30 mm        |
| Front/Middle/Rear roof arch  | 5–15 mm         |
| Roof-side rail               |                 |
| Front/Middle/Rear joint      | 5–15 mm         |

After the basic simplified FE model was configured, the diameters of ten parts were set as design variables, as shown in Table 2. A total of 2000 models were created by changing the design variables; these models were used for analysis. In order to extract variables without any overlap in the design space during this process, the Latin hypercube sampling technique was used. In Latin hypercube sampling, each probability variable is divided into N ranges so that samples can be extracted from the overall sample space considered. Then, one sample is extracted from each interval so that N samples are extracted without overlap. It is a sampling technique that can consider all design spaces without repetition [22–24].

#### 2.2. Data Structure

(a) Simplified finite element (FE) roof crush model from 2012 Toyota Camry model [21], (b) Oblique collision to simulate the rollover-driven roof impact and boundary conditions.
Hierarchical clustering has excellent consistency and reproducibility in the clustering process, and it has the advantage of being able to perform clustering even if the user does not set the number of clusters beforehand. As shown in Figure 3a, a random number of clusters can be set based on the displacement data points of each model [13,14].

In this study, agglomerative (bottom-up) hierarchical clustering algorithms were used to construct a cluster hierarchy that is commonly displayed as a tree diagram called a dendrogram [13,14]. Hierarchical clustering has excellent consistency and reproducibility in the clustering process, and it has the advantage of being able to perform clustering even if the user does not set the number of clusters beforehand. As shown in Figure 3a, a random number of clusters can be set based on the displacement data of the 2000 BIW models obtained in the above analysis results. The vertical axis of the dendrogram represents the distance or dissimilarity between clusters. The horizontal axis represents the case number of the 2000 model clusters. The model initially has each object in a separate cluster. At each step, two clusters that are the most similar are combined to form a single new cluster. Here, the Minkowski distance was used to quantify the dissimilarity from the displacement datapoints of each model [13,14]. The Minkowski distance between the \(i\)-th and \(j\)-th BIW models is defined as

\[
d_{ij} = \sqrt[p]{\sum_{k=1}^{2436} |r_{ik} - r_{jk}|^p}.
\]

Here, \(r_{ij}\) is the components of the displacement data vector in Table 3. Additionally, \(i\) and \(j\) are the case number of the 2000 BIW models \((i, j = 1, 2, \ldots, 2000)\). Here, \(p\) is a user-defined exponent related to what can be called a curvature bias. The Minkowski distance is a normed vector space, which can be considered to be a generalization of both the Euclidean distance \((p = 2)\) and the Manhattan distance \((p = 1)\). We calculated the Minkowski distances for the cases of \(p = 1.5, 2,\) and \(2.5\) to compare cluster distributions. Finally, the BIW models were classified, and the resultant dendrogram was plotted in Figure 3a based on the distance function calculated at \(p = 2\). In this study, all 2000 models were separated into clusters of three quantities (10, 15, and 20), as shown in Figure 3b.
Figure 3. (a) A cluster hierarchy (dendrogram) of 2000 BIW models, (b) split lines for the three cases of clustering (10, 15, and 20), and (c) simple diagram of the hierarchical clustering procedure.

The hierarchical clustering process of the 2000 BIW models is summarized in Figure 3c. At each step, the two clusters with the shortest distance are combined to form a single new cluster. Afterward, the distances between all the clusters created by the above process are recalculated, and the process of combining two clusters with the shortest distance is repeated until ultimately there is a single cluster.

Table 4 summarizes the number of models, reaction force, and mass in each cluster for the three cases of clusters (10, 15, and 20). Figure 4 shows the average reaction force of 20 clusters with error bars at each cluster. The error bars represent the standard deviation (SD) of the data. Strong clusters are defined as those with high mean reaction forces. Then, we selected only the strong clusters for optimization by eliminating weak clusters with low forces from the obtained clusters. An RSM was constructed for optimization using the selected strong clusters. It is noted that the proposed pattern recognition technique prevented design variables of weak clusters from being in the optimal solution during the optimization process.
Table 4. Three types of clustering obtained from 2000 BIW models (a) 10 clusters; (b) 15 clusters; and (c) 20 clusters.

| Cluster | N | Reaction force (kN) | Mass (kg) |
|---------|---|---------------------|-----------|
|         |   | Mean    | SD       | Mean    | SD       |
|         |   | (a) 10 clusters |          | (b) 15 clusters |          | (c) 20 clusters |          |
| 1       | 31| 42.226  | 16.727   | 25.047  | 3.723    |
| 2       | 45| 43.228  | 16.645   | 25.119  | 5.639    |
| 3       | 193| 44.570  | 18.525   | 25.265  | 5.691    |
| 4       | 169| 45.150  | 18.200   | 26.368  | 5.095    |
| 5       | 314| 45.560  | 20.411   | 26.011  | 5.298    |
| 6       | 383| 45.606  | 18.219   | 26.188  | 5.520    |
| 7       | 134| 45.706  | 17.067   | 26.526  | 5.728    |
| 8       | 286| 46.019  | 17.466   | 26.604  | 5.275    |
| 9       | 265| 46.475  | 17.027   | 27.205  | 5.413    |
| 10      | 83 | 48.079  | 18.880   | 26.116  | 5.473    |
|         | Total 2000 | 45.339  | 18.234     | 26.305  | 5.405      |
|         | 1   | 100     | 47.379    | 18.178   | 26.810   | 6.048    |
| 2       | 23 | 45.892  | 13.661   | 25.771  | 4.707    |
| 3       | 114| 43.118  | 17.645   | 24.964  | 5.956    |
| 4       | 31 | 42.226  | 16.727   | 25.047  | 3.723    |
| 5       | 83 | 48.079  | 18.880   | 26.116  | 5.473    |
| 6       | 45 | 43.228  | 16.645   | 25.119  | 5.639    |
| 7       | 79 | 46.667  | 19.649   | 25.699  | 5.887    |
| 8       | 265| 46.475  | 17.027   | 27.205  | 5.413    |
| 9       | 31 | 40.306  | 11.499   | 25.610  | 4.508    |
| 10      | 74 | 49.502  | 17.422   | 27.692  | 5.460    |
| 11      | 189| 44.670  | 17.781   | 26.279  | 5.230    |
| 12      | 69 | 44.285  | 17.140   | 25.527  | 5.174    |
| 13      | 269| 45.150  | 18.200   | 26.368  | 5.095    |
| 14      | 314| 45.997  | 18.461   | 26.333  | 5.591    |
| 15      | 314| 45.560  | 20.411   | 26.011  | 5.298    |
|         | Total 2000 | 45.339  | 18.234     | 26.305  | 5.405      |
|         | 1   | 14      | 40.111    | 13.056   | 27.193   | 5.390    |
| 2       | 31 | 37.765  | 71.632   | 26.154  | 4.229    |
| 3       | 74 | 47.038  | 21.681   | 26.096  | 4.944    |
| 4       | 115| 48.405  | 21.480   | 26.273  | 6.043    |
| 5       | 166| 48.194  | 13.619   | 25.191  | 5.192    |
| 6       | 69 | 40.703  | 18.265   | 25.988  | 5.356    |
| 7       | 11 | 52.670  | 18.904   | 27.491  | 5.292    |
| 8       | 20 | 51.320  | 23.562   | 29.554  | 5.324    |
| 9       | 88 | 48.875  | 13.453   | 28.712  | 4.498    |
| 10      | 226| 47.055  | 15.403   | 27.695  | 5.163    |
| 11      | 23 | 13.380  | 13.847   | 24.469  | 4.109    |
| 12      | 74 | 50.268  | 19.279   | 23.300  | 4.611    |
| 13      | 79 | 40.165  | 15.901   | 25.687  | 5.354    |
| 14      | 114| 39.066  | 12.234   | 27.969  | 4.938    |
| 15      | 31 | 46.426  | 17.428   | 22.834  | 4.948    |
| 16      | 100| 41.189  | 12.920   | 25.352  | 4.548    |
| 17      | 69 | 54.635  | 18.792   | 26.042  | 4.055    |
| 18      | 314| 48.184  | 19.037   | 25.728  | 5.386    |
| 19      | 83 | 46.130  | 11.638   | 28.735  | 4.186    |
| 20      | 269| 50.176  | 22.698   | 25.637  | 6.375    |
|         | Total 2000 | 43.339  | 18.234     | 26.305  | 5.405      |
produced 10 clusters as output, 0.1% in the trained net that produced 15 clusters as output, and 3.5% in the trained net that produced 20 clusters as output. Therefore, nets with excellent performance could be created.

A pattern consists of a number of features (also known as attributes). The classification accuracy of a classifier is judged by the number of testing patterns it has classified correctly. There are several approaches in pattern recognition; this study used a pattern recognition approach that employed a neural network [10,26]. When pattern recognition is performed through a neural network, required prior knowledge can be reduced. If there are sufficient quantities and layers of neurons, it has the advantage of theoretically being able to create any complex network and easy to use compared with other techniques. Lastly, it has high expected performance values.

It is noted that neural networks themselves can lead to many different classifiers depending on how they are trained, involving the danger of overtraining the classifier as the number of network parameters increases. Therefore, the regularization of neural networks may be necessary [10]. In this study, the Bayesian regularization technique was used as a training algorithm. The Bayesian algorithm is a technique that updates the weights and biases to minimize linear combinations of weights and squared errors according to Levenberg–Marquardt optimization. It is known as a combination of the Gauss–Newton technique and the gradient descent technique, which improves upon the existing Gauss–Newton technique [27,28].

A total of 100 hidden layers were used for training. The input data were the 10 design variables of all 2000 models, and the input data were 10, 15, and 20 clusters obtained in the above clustering results. A total of three rounds of training were performed. In order to intuitively show the error after training, the predicted output when all input data were entered in the completely trained net was compared with the cluster output obtained in the clustering results, and the difference was defined as a percent error to confirm the performance of the net. The percent error was 0.05% in the trained net that produced 10 clusters as output, 0.1% in the trained net that produced 15 clusters as output, and 3.5% in the trained net that produced 20 clusters as output. Therefore, nets with excellent performance could be created.

4. Optimization

4.1. Proposed Optimization Technique

Unlike the RSM configuration in traditional optimization techniques, this study proposed a new optimization technique by employing a pattern recognition technique based on the clusters obtained.
as the results of the above clustering process. Table 5 summarizes the optimization procedures of the two methods. The second-order polynomial model was used as a means of making the RSM of mass and reaction force functions with 10 design variables as follows

\[ y = a_0 + \sum_{i=1}^{10} a_i x_i + \sum_{i<j}^{10} a_{ij} x_i x_j \]  

(2)

Here, \( y \) is the output value (reaction force or mass) for the optimization and \( x_i \) is the 10 design variables of the BIW models shown in Tables 2 and 3. The numbers of the coefficients \( a_i \) and \( a_{ij} \) of the first-order and the second-order terms were 10 and 55, respectively. The constant \( a_0 \) and the coefficients were determined through the linear regression method that satisfies the R-square value over 0.9. The genetic algorithm in MATLAB software was used to implement the optimization for the RSM configuration.

**Table 5.** Comparison of traditional and proposed optimization techniques.

| Traditional Optimization Method |
|---------------------------------|
| (1) Configure RSM using all 2000 models |
| (2) Perform optimization under the following conditions: |
| Objective function: Maximizing reaction force |
| Constraints: Targeting mass (26, 27, 28, 29, 30 kg) |

| Proposed Optimization Method |
|------------------------------|
| (1) Configure RSM using only \( N \) number of strong clusters obtained as clustering results |
| (2) Perform optimization under the following conditions: |
| Objective function: Maximizing reaction force |
| Constraints: Targeting mass (26, 27, 28, 29, 30 kg) & avoiding \( N \) weak clusters |

This study enacted constraints for robust design to ensure that the optimization results did not belong to the weak clusters with low reaction forces. To find the optimum value set that maximizes reaction force at the specific targeting mass, the RSM of reaction force was set as an objective, whereas the mass was treated as an equality constraint. Then, the recognized neural network worked as an inequality constraint by including only strong clusters with large reaction forces for the three clustering cases in Table 4. The constraint tolerance was set to be 0.001. After finding clusters for the input variable set, the neural network rearranged these clusters in an ascending order to simplify constraints in optimization. The pseudocode of the optimizing process, that is, the description of the operating principle of the optimization algorithm, is shown in the Appendix A.

### 4.2. Optimization Results

The RSM was configured according to the traditional and proposed optimization methods listed in Table 5, and a genetic algorithm was used to optimize each case. The genetic algorithm has the disadvantage of requiring a lot of computation, but it has the advantage of not depending on an initial design point. Thus, there is little concern over it converging at a local optimum point, and parallel computation is possible. It also has the advantage of being applicable to almost all variable types as well as discontinuous problems. The predicted response obtained from the optimization results was compared with the actual response obtained from performing FEA on the model composed of optimized design variables obtained during optimization. Table 6 summarizes the optimization results for the case of a mass of 30 kg using 5 strong clusters from 15 clustering type. In this case,
757 models were used for the RSM to calculate the optimized design variables and the corresponding reaction forces.

Table 6. Optimal design values and maximal reaction forces by traditional and proposed optimizations for the targeting mass of 30 kg using 5 strong clusters from 15 clustering types (case 6 in Table 7).

|                     | Traditional Optimization                                      | Proposed Optimization          |
|---------------------|--------------------------------------------------------------|---------------------------------|
| Approximation       | R squared mass: 0.9999                                      | R squared mass: 0.9999          |
|                     | R squared reaction force: 0.924                             | R squared reaction force: 0.9717|
| Optimization results| Reaction force: 107.5 kN Mass: 30.00 kg                     | Reaction force: 119.4 kN Mass: 30.00 kg |
| Actual results      | Reaction force: 116.0 kN Mass: 30.00 kg                     | Reaction force: 123.3 kN Mass: 30.00 kg |
| Optimal design variables | Center pillar: 5.00 mm Side rail: 8.42 mm  | Center pillar: 5.00 mm Side rail: 9.49 mm |
|                     | Front pillar: 15.00 mm Mid joint: 3.36 mm                  | Front pillar: 15.00 mm Mid joint: 2.50 mm |
|                     | Middle roof arch: 5.00 mm Rear pillar: 14.6 mm             | Middle roof arch: 5.00 mm Rear pillar: 14.5 mm |
|                     | Rear joint: 7.50 mm Front joint: 7.50 mm                   | Rear joint: 7.50 mm Front joint: 7.50 mm |
|                     | Rear roof arch: 13.6 mm Front roof arch: 15.00 mm          | Rear roof arch: 12.2 mm Front roof arch: 15.00 mm |

In this study, all 2000 BIW models were clustered into three quantities of clusters: 10, 15, and 20, as shown in Table 4. Based on the optimization process in Table 5, the proposed optimization results were summarized in Table 7 for different clustering types (3, 5, and 7 strong clusters in the 10 clustering; 3, 5, 7, 8, 10, and 12 strong clusters in the 15 clustering; 3, 5, 7, 13, 15, and 17 strong clusters in the 20 clustering) with the 5 cases of targeting mass (26, 27, 28, 29, and 30 kg). In most cases, the maximal reaction forces obtained by the proposed optimization process were improved compared with the values by the traditional optimization. Here, OM values represent the calculated (predicted) results of the reaction force obtained using the proposed optimization method. Additionally, FEA values mean the actual reaction force calculated by a commercial finite element software, ABAQUS, using the optimal design variables obtained by the optimization process. The error rate is calculated by \(|\text{OM} – \text{FEA}| / \text{OM} (%)\).

Figure 5 shows the improvements in the proposed optimization results for each cluster compared with the traditional optimization. Compared with the traditional method, when the RSM configuration used three strong clusters that were obtained from the 10 clusters (case 2 in Table 7), the overall improvement was 3.55%. When the RSM configuration used five strong clusters that were obtained from the 15 clusters (case 6 in Table 7), the overall improvement was 3.48%. When the RSM configuration used seven strong clusters that were obtained from the 20 clusters (case 13 in Table 7), the overall improvement was 4.64%. In the remaining cases, a small range of performance improvement was also confirmed.
models belonging to strong clusters, the expected performance values were high, and it was possible to

Figure 5. Improvement rate of the maximal reaction force by the proposed optimization compared with that of the traditional optimization for the three cluster types (cases 2, 6, and 13 in Table 7).

Table 7. Reaction forces and error rate by the proposed optimization method (OM) in various clustering cases.

| Case | Cluster Type            | Model Numbers | Targeting Mass 26 kg | Targeting Mass 27 kg | Targeting Mass 28 kg | Targeting Mass 29 kg | Targeting Mass 30 kg |
|------|------------------------|---------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 1    | Traditional optimization | 2000          | OM (FEA [kN]) 89.95/97.71 Error rate (%) 8.63 | OM (FEA [kN]) 94.62/101.7 Error rate (%) 7.48 | OM (FEA [kN]) 99.13/107.1 Error rate (%) 8.04 | OM (FEA [kN]) 103.4/110.8 Error rate (%) 7.16 | OM (FEA [kN]) 107.5/116.0 Error rate (%) 7.91 |
| 2    | 3 strong clusters from 10 clustering | 603           | 101.9/100.4 1.47 | 109.6/103.9 5.03 | 112.9/109.7 2.83 | 116.2/116.0 0.17 | 119.4/122.6 2.68 |
| 3    | 5 strong clusters from 10 clustering | 1283          | 95.84/96.23 0.44 | 100.3/102.0 1.69 | 104.6/108.0 3.25 | 108.7/113.2 4.14 | 112.6/117.5 5.06 |
| 4    | 7 strong clusters from 10 clustering | 1497          | 94.83/97.26 2.56 | 99.50/102.6 3.12 | 104.0/108.1 3.95 | 108.3/112.9 4.29 | 112.3/117.5 4.59 |
| 5    | 3 strong clusters from 15 clustering | 174           | 112.3/74.55 3.36 | 119.2/81.23 3.19 | 126.0/64.75 3.27 | 132.9/97.30 3.43 | 139.7/90.88 3.49 |
| 6    | 5 strong clusters from 15 clustering | 757           | 103.8/99.38 4.23 | 108.0/102.8 4.81 | 111.9/110.0 3.27 | 115.7/117.0 3.43 | 119.4/123.3 3.27 |
| 7    | 7 strong clusters from 15 clustering | 1260          | 95.70/95.94 0.25 | 100.1/102.4 2.30 | 104.3/108.1 3.64 | 108.3/113.4 4.71 | 112.2/118.6 5.79 |
| 8    | 8 strong clusters from 15 clustering | 1291          | 95.49/96.23 0.77 | 99.89/102.3 2.41 | 101.4/107.9 3.65 | 108.1/113.4 4.9 | 112.0/119.0 6.25 |
| 9    | 10 strong clusters from 15 clustering | 1397         | 95.66/96.02 0.38 | 100.1/102.5 2.40 | 104.3/107.9 3.45 | 108.3/113.5 4.80 | 112.2/118.9 5.97 |
| 10   | 12 strong clusters from 15 clustering | 1576         | 93.37/99.00 6.03 | 97.97/103.1 5.24 | 102.6/107.9 5.17 | 107.0/113.2 5.79 | 111.1/118.4 6.57 |
| 11   | 3 strong clusters from 20 clustering | 100          | 147.5/56.49 61.7 | 156.7/59.47 62.0 | 167.0/56.82 65.9 | 177.8/56.12 68.4 | 188.6/61.16 67.7 |
| 12   | 5 strong clusters from 20 clustering | 443          | 101.5/101.3 0.20 | 106.4/107.1 0.66 | 111.3/111.4 0.09 | 116.2/116.0 0.17 | 120.9/121.4 0.41 |
| 13   | 7 strong clusters from 20 clustering | 646          | 100.8/107.0 0.79 | 105.8/107.2 1.32 | 110.7/111.4 0.63 | 115.6/116.3 0.61 | 120.3/121.6 1.08 |
| 14   | 13 strong clusters from 20 clustering | 1397       | 95.61/96.07 0.48 | 100.1/102.6 2.50 | 104.3/108.3 3.55 | 108.3/111.9 4.71 | 112.2/118.8 5.88 |
| 15   | 15 strong clusters from 20 clustering | 1596       | 94.86/96.99 2.25 | 99.56/102.0 2.45 | 104.0/108.3 3.55 | 108.3/111.9 4.71 | 112.4/118.1 5.07 |
| 16   | 17 strong clusters from 20 clustering | 1689       | 93.20/97.60 4.61 | 97.98/102.9 5.02 | 102.6/107.5 4.78 | 106.9/112.6 5.33 | 111.1/117.6 5.85 |

Figure 6 shows the error rates between the optimization results and the FEA results according to the model sample quantities, which are shown in Table 7. When the RSM was configured using only the models belonging to strong clusters, the expected performance values were high, and it was possible to
perform better optimization than before. However, when the RMS was configured using less than 10% of the overall 2000 models in strong clusters, the optimization prediction performance was not reliable. In a comparison of the optimization and the FEA results, 100 (5% of the total) models were used when three strong clusters were selected from the 20 clusters to perform the RMS configuration (case 11 in Table 7), and 174 models (8.7% of the total) were used when three strong clusters were selected from the 15 clusters to perform RMS configuration (case 5 in Table 7). In these cases, an error rate of 30% or more occurred. This appears to be the occurrence of an underfitting phenomenon caused by the use of too few models in the RMS configuration. The results of using all 2000 models to validate the above response surfaces of the two models confirmed that the strong models, which were aimed at configuring the RSM, were able to perform modeling sufficiently with an R-squared of 0.9. However, the design of all the models was very inadequate with an R-squared of less than 0.2.

![Figure 6. Error rates by the number of samples of various clustering types.](image)

When over 20% of all the models were used to configure the RSM, the error rate was similar to or less than that of the traditional optimization technique, and the optimization performance was excellent compared with that of the traditional technique. Therefore, when the RMS was configured using more than 20% of all models, optimization reliability and safety were guaranteed. The clustering results of the area shown by the dotted line in Figure 6 were excellent cases in terms of error rate. The cases in which the number of models was 20%–30% showed better results than those in which the number of models was large.

5. Conclusions

This study proposed a new optimization method that adds artificial intelligence techniques, such as clustering and pattern recognition, to traditional optimization techniques in an optimal vehicle BIW design that considers roof strength. The optimization results of the proposed technique were compared with the results obtained from the existing technique to verify their excellence. To this end, clustering and pattern recognition training with 10, 15, and 20 clusters was performed on 2000 simplified BIW analysis models, and the following conclusions were drawn.

1. When the RSM was configured using only models that belonged to strong clusters, it was expected to exhibit an improved performance compared with the existing technique, because it was configured only with items with excellent performance. As the number of strong clusters configuring the RSM decreased, the expected optimal performance became excellent, and as more models were included in the RSM configuration, the performance deteriorated. For the
RSM composed of all 2000 models, the optimization results show that the predicted value of the reaction force was 89.95, 94.62, 99.13, 103.4, and 107.5 kN for the five cases of targeting mass (26, 27, 28, 29, and 30 kg), respectively. However, in the optimization results derived from the other RSM configured by considering only strong clusters, the predictions exceeded all of the optimization performance values of the traditional technique.

2) When the RSM was configured with only the models belonging to strong clusters, the expected performance values were high, and it was possible to perform better optimization than that of the existing technique. However, when less than 10% of the overall 2000 models were selected in strong clusters for the RSM configuration (cases 5 and 11 in Table 7), the optimized prediction performance was not reliable.

3) In the other results, for which more than 20% of the total models were used to configure the RSM, a similar or smaller error rate than that of the traditional optimization technique was observed, and the optimization performance was also better than that of the traditional technique. Therefore, it was considered safe to use more than 20% of all the models to configure the RSM in order to ensure optimization reliability.

Author Contributions: Y.-S.S. analyzed clustering and pattern recognition. H.-J.L. contributed to the modeling of the roof crush. Y.-J.C. and S.-Y.L. contributed to the optimization analysis.

Funding: This research was supported by the Nuclear Research and Development Program (20161520101210) of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) funded by the Ministry of Trade and Industry and the National Research Foundation of Korea grant (NRF-2018M2A8A4083765) funded by the Ministry of Science ICT.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Algorithm A1: Pseudocode of Optimization Algorithm (Targeting 20 kg and 5 strong clusters from 15 clustering)

| Algorithm A1 |
|--------------|
| **Input:** |
| \( F \) — Reaction force RSM |
| \( f \) — Mass RSM function |
| \( \text{Net} \) — Pretrained Network |
| \( S \) — Design variable boundary |
| **Output:** |
| Optimum \( X \) (10 dimension) for maximum reaction force |
| **Function** \( [x] = \text{GA-opt}(F, nvars, C_{ineq}, C_{eq}, S) \) |
| \( \text{Netfcn} \) — transform cluster vectors to indices |
| \( nvars \) = dimension of variables set, 10 |
| \( C_{ineq} = (5 - \text{Netfcn}(X_t)) \) // Using 5 strong clusters |
| \( C_{eq} = (f-30) \) // Targeting mass 30 kg |
| \( t \leftarrow 0 \) |
| Initialize \( X_t \) to random individuals from \( S \) |
| **While** termination condition not met |
| do |
| Select elements from \( X_t \) |
| Recombine and mutate variables |
| Evaluate Fitness-GA (S, modified variables) |
| \( X_{t+1} \) (updated variables set) |
| \( t \leftarrow t+1 \) |
| **Return** best \( X_t \) maximizes \( F \) |
| **End** |
References

1. Nicholas, L. Review of Technical Literature and Trends Related to Automobile Mass-Reduction Technology; Report UCD-ITS-RR-10-10; Institute of Transportation Studies: Davis, CA, USA, 2010.
2. Joost, W. Reducing vehicle weight and improving U.S. energy efficiency using integrated computational materials engineering. *JOM* 2012, 64, 1032–1038. [CrossRef]
3. Martin, G.; Marc, S.; Lukas, R.; Gundolf, K.; Elmar, B. Super light car—Lightweight construction thanks to a multi-material design and function integration. *Eur. Transp. Res. Rev.* 2009, 1, 5–10.
4. Mundo, D.; Hadjit, R.; Donders, S.; Brughmans, M.; Mas, P.; Desmet, W. Simplified modeling of joints and beam-like structures for BIW optimization in a concept phase of the vehicle design process. *Finite Elem. Anal. Des.* 2009, 45, 456–462. [CrossRef]
5. Wang, Z. Survey of optimization methods for BIW lightweight design. *MATEC Web Conf.* 2017, 128, 050002.
6. Sreekanta, D.; Sudip, B.; Pratunu, G. Roof Strength Requirement for Vehicles Involved in Rollover Crash; SAE Technical Paper, 2008-01-0510; Society of Automotive Engineers: Warrendale, PA, USA, 2008.
7. Ryu, S.U.; Wook, J.; Shin, Y.C.; Bae, S.K.; Lee, K.B. Study of optimal body structural design for coupe-type vehicles in rollover events. In Proceedings of the International Conference on the Enhanced Safety of Vehicles, Seoul, Korea, 27–30 May 2013; Volume 2, pp. 781–788.
8. Christensen, J.; Bastien, C.; Blundell, M.V. Effects of roof crush loading scenario upon body in white using topology optimization. *Int. J. Crashworthiness* 2012, 17, 29–38. [CrossRef]
9. Ichiro, S.; Daisei, A. Pattern recognition and machine learning approach for BIW deformation mode control. In Proceedings of the JSAE Annual Spring Conference, Yokohama, Japan, 21–23 May 2014.
10. Jain, A.K.; Duin, R.P.W.; Mao, J. Statistical pattern recognition: A review. *IEEE Trans. Pattern Anal.* 2000, 22, 4–37. [CrossRef]
11. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, perspectives and prospects. *Science* 2015, 349, 255–260. [CrossRef]
12. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaaadi, F.E. A survey of deep neural network architectures and their applications. *Neurocomputing* 2017, 234, 11–26. [CrossRef]
13. Xu, D.; Tian, Y.A. Comprehensive survey of clustering algorithms. *Ann. Data Sci.* 2015, 2, 165–193. [CrossRef]
14. Saxena, A.; Prasad, M.; Gupta, A.; Bharill, N.; Patel, O.P.; Tiwari, A.; Lin, C.T. A review of clustering techniques and developments. *Neurocomputing* 2017, 267, 664–681. [CrossRef]
15. Mukhtar, A.; Likun, X. Vehicle detection techniques for collision avoidance systems: A review. *IEEE Trans. Intell. Syst. Transp.* 2015, 16, 2318–2338. [CrossRef]
16. Carrio, A.; Sampedro, C.; Rodriguez-Ramos, A. A review of deep learning methods for unmanned aerial vehicles. *J. Sens.* 2017, 2, 1–13. [CrossRef]
17. Liu, R.; Yang, B.; Zio, E.; Chen, X. Artificial intelligence for fault diagnosis of rotating machinery: A review. *Mech. Syst. Signal Process.* 2018, 108, 33–47. [CrossRef]
18. Reza, M.; Sina, S.H.; Sami, S.H.; Mohammad, A. Performance of circular saw machine using imperialist competitive algorithm and fuzzy clustering technique. *Neural Comput. Appl.* 2018, 29, 283–292.
19. Yaghoobi, A.; Bakhshi-Jooybari, M.; Gorji, A.; Baseri, H. Application of adaptive neuro fuzzy inference system and genetic algorithm for pressure path optimization in sheet hydroforming process. *Int. J. Adv. Manuf. Technol.* 2016, 86, 2667–2677. [CrossRef]
20. Vogl, G.W.; Weiss, B.A.; Helu, M. A review of diagnostic and prognostic capabilities and best practices for manufacturing. *J. Intell. Manuf.* 2019, 30, 79–95. [CrossRef] [PubMed]
21. LS-DYNA Examples. 2012 Toyota Camry Finite Element Model Version. Available online: https://www.dynaexamples.com/implicit/implicit_roofcrush (accessed on 5 September 2018).
22. Mckay, M.D.; Beckman, R.J.; Conover, W.J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 2000, 42, 55–61. [CrossRef]
23. Shields, M.D.; Zhang, J. The generalization of Latin hypercube sampling. *Reliab. Eng. Syst. Saf.* 2016, 148, 96–108. [CrossRef]
24. Ruiyi, S.; Liangjin, G.; Zijie, F. Multi-objective optimization for bus body with strength and rollover safety constraints based on surrogate models. *Struct. Multidiscip. Optim.* 2011, 44, 431–441.
25. Seo, S.I.; Lee, M.K.; Kim, S.J.; Kim, N.S. Robust optimum design of a bearingless rotation motor using the kriging model. *Int. J. Precis. Eng. Man* 2011, 12, 1043–1050. [CrossRef]
26. Theodoridis, S.; Pikrakis, A.; Koutrombas, K.; Cavouras, D. *Introduction to Pattern Recognition: A Matlab Approach*; Elsevier: Amsterdam, The Netherlands, 2010.

27. Garg, V.K.; Bansal, R.K. Analysis of various computing techniques for diagnosis of sleep disorders. *Int. J. Eng. Res. Comput. Sci. Eng.* 2018, 5, 84–90.

28. Bhavan, S.; Venugopalan, K. Comparison of neural network training functions for Hematoma classification in brain CT images. *IOSR J. Comput. Eng.* 2014, 16, 31–35.

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).