ACS APPLIED MATERIALS & INTERFACES

www.acsami.org

Research Article

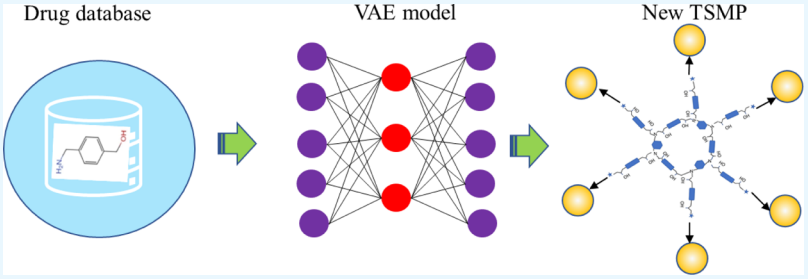
From Drug Molecules to Thermoset Shape Memory Polymers: A Machine Learning Approach

Cheng Yan, Xiaming Feng, and Guoqiang Li*

Cite This: *ACS Appl. Mater. Interfaces* 2021, 13, 60508–60521

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information



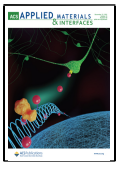
ABSTRACT: Ultraviolet (UV)-curable thermoset shape memory polymers (TSMPs) with high recovery stress but mild glass transition temperature (T_g) are highly desired for 3D/4D printing lightweight load-bearing structures and devices. However, a bottleneck is that high recovery stress usually means high T_g. For a few TSMPs with high recovery stress, their T_g values are close to the decomposition temperature, and thus, the shape memory effect cannot be triggered safely and effectively. While machine learning (ML) has served as a useful tool to discover new materials and drugs, the grand challenge of using ML to discover new TSMPs persists in the very limited data available. Here, we report an enhanced ML approach by combining the transfer learning–variational autoencoder with a weighted-vector combination method. By learning a large data set with drug molecules in a pretraining process, we were able to effectively map the TSMPs to a hidden space that is much closer to a Gaussian distribution. Through this approach, we created a large compositional space and were able to discover five new types of UV-curable TSMPs with desired properties, one of which was validated by the experiments. Our contribution includes (1) representing the features of TSMPs by drug molecules to overcome the barrier of a limited training data set and (2) developing a ML framework that is able to overcome the barrier of mapping the molar ratio information. It is shown that this approach can effectively learn TSMP features by utilizing the relatedness between the data-scarce (and biased) TSMP target and data-abundant drug source, and the result is much more accurate and more robust than the benchmark set by the support vector machine method using direct label encoding and Morgan encoding. Therefore, it is believed that this framework is a state-of-the-art study in the TSMP field. This study opens new opportunities for discovering not only new TSMPs but also other thermoset polymers.

KEYWORDS: *machine learning, variational autoencoder, transfer learning, shape memory polymer, material discovery, drug molecules, 4D printing*

1. INTRODUCTION

In the past decade, with the development of hardware[1] and new algorithms,[2−4] machine learning (ML) has become capable of learning more hidden features from targets and hence has been becoming a more and more popular tool in many research fields.[5−8] In the field of materials science, as indicated by Yan et al.,[9] ML bears three prominent advantages, that is, (a) it overcomes the time-consuming *trial-and-error* approach in traditional methods,[10−12] (b) it is at least tens or hundreds of times faster than the traditional molecular or atomistic computational models,[13,14] and (c) it bears a universality as compared with the traditional solid mechanics models.[15−19] As such, ML has also been widely adopted for a variety of materials, especially for drugs. Researchers have established more than 10 databases for drugs. For example, PDBbind[20] provides the receptor–ligand binding data for resolved protein structures, Pubchem[21] supplies a wide range of chemical information involving physical–chemical properties and biomolecular interaction, Uniprot[22] involves a large amount of data for protein sequence homology and protein ID retrieving, RCSB PDB[23] provides information for the protein 3D structure, and so forth, wherein tens of millions of molecules can be found. By utilizing these databases,

Received: October 29, 2021
Accepted: November 18, 2021
Published: December 8, 2021



https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508–60521

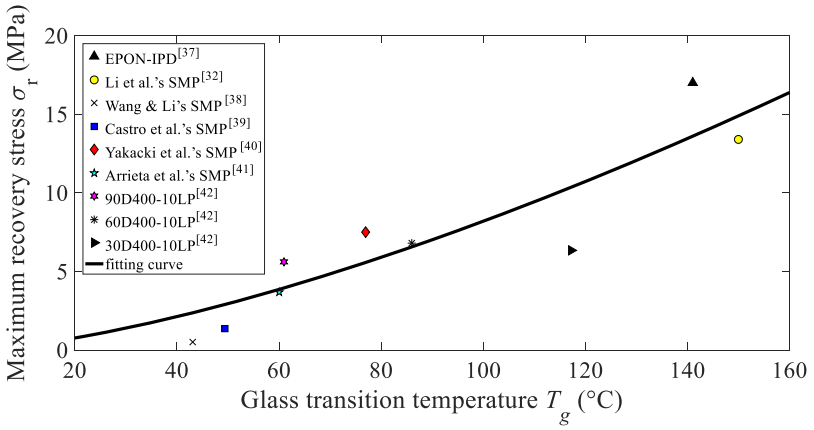


Figure 1. Approximate relation between T_g and maximum recovery stress for some SMPs.[32,37−42]

investigators have already achieved a lot of success. For example, through Pubchem, Butkiewicz et al. assembled multiple data sets to establish a new benchmark platform for drug target proteins;[24] Zhang et al. implemented data points from PDBbind to train a new deep neural network model, which is able to predict the binding affinity of a protein−ligand complex;[25] and based on PubMed and ClinicalTrials.gov databases, Ong et al. leveraged Vaxign and the newly developed ML-based Vaxign−ML reverse vaccinology tools to predict COVID-19 vaccine candidates.[26]

Meanwhile, in the past decades, due to the 3D printability, good mechanical properties, and an excellent shape memory effect (SME), ultraviolet (UV)-curable thermoset shape memory polymers (TSMPs) have found wide applications in actuators,[27] proppants,[28] flame retardancy,[29] sealants,[30] self-healing,[31,32] and many more.[33] By using UV-curable TSMPs and 3D printing[34] via digital light processing or direct ink writing, structures with complex shapes and functionalities can be manufactured such as deployable structures[35] and soft robotics.[36] For these applications, they usually require higher recovery stress, for example, using the UV-curable TSMP for manufacturing a multifunctional lightweight load-carrying architecture.[34] However, most UV-curable TSMPs with high strength, high stiffness, and high recovery stress also have high T_g. Taking T_g as an example, as shown in Figure 1, the maximum recovery stress for SMPs is proportional to T_g. Sometimes, very high T_g is a barrier because the UV-curable TSMP must be heated up significantly to trigger the SME, which may hinder the applications because the high trigger temperature may be close to the decomposition temperature of the polymer. Therefore, designing new TSMPs with high recovery stress but moderate T_g is highly desired. However, due to the conflict requirement between moderate T_g, which prefers softer UV-curable TSMPs with mobile molecular segments, and high recovery stress, which prefers a stiffer network with less mobile molecular segments, it is a grand challenge for human intelligence to balance this conflict requirement; ML may be a useful tool.

Unfortunately, very few ML models are available for discovering new TSMPs[9,43] due to three main difficulties, that is, a complex polymer network, the lack of a public TSMP database, and multifactor-dependent experimental results. To our knowledge, the first and only ML framework for new TSMP discovery was presented by Yan et al.,[9] wherein they established a dual-convolutional neural network (CNN) that is able to perform a forward design by predicting the T_g and recovery stress. Although this model can discover some new TSMPs with higher recovery stress than the known data available by then, it bears three limitations. First, the accuracy of the model needs to be improved. To be specific, the mean average percentage error (MAPE) of the models can just hit 28% and hence leaves a large room for improvement. Second, the data set in that study is small, which could lead to overfitting. As indicated by Yan et al.,[9] only about 100 data sets were collected for the training, which could make the neural network difficult to fit the appropriate values for a large number of parameters (weights) in the CNN layers. This is partially proved by the gap of the loss between the training data and test data, which is more than 20%, thus empirically an overfitting may exist there. Last but not the least, the model did not take the molar ratio into consideration. In that study, they leveraged BigSMILES[44] for fingerprinting the polymers, which deals with the bond connectivity but is not able to provide the information for the molar ratio. This makes the ML model ignore an essential feature of SMPs. This is because the molar ratio determines how much of each reagent is actually involved in the reaction and what is the final topology of the network and thus plays a critical role in polymer properties. For instance, for the SMP material synthesized by di(ethylene glycol)dimethacrylate (DEGDMA) and *tert*-butyl acrylate, Yakacki et al.[40] changed the molar amount of DEGDMA from 0.025 to 0.134 mol, which then significantly improves the rubbery modulus (E_r) by 6.67 times (from 1.5 to 11.51 MPa); in another SMP network, Barszczewska-Rybarek et al.[45] changed the molar ratio between methyl methacrylate and triethylene glycol dimethacrylate from 0.01:0.99 to 0.20:0.80 and found that the rubbery modulus was increased by about 14-fold (from 2.08 to 30.89 MPa).

For the first and second difficulties, we can attribute them to the inadequate training data and the use of an unadvanced ML approach. In fact, there exists a big gap between the need for TSMP data and the actual available TSMP data. On one hand, the features of TSMP networks are various and sophisticated, that is, the complete information on the multiple length scale structures of the cross-linked TSMP networks encompasses atomistic, topological, and morphological structures, and hence make an ML model extremely hungry for data. On the other hand, TSMPs belongs to a new branch of smart material, and thus, not enough data can be found in the literature. For example, only 6198 papers can be found in the largest accessible citation database "Web of Science" involving the keyword "shape memory polymer", which implies that, at most,

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521

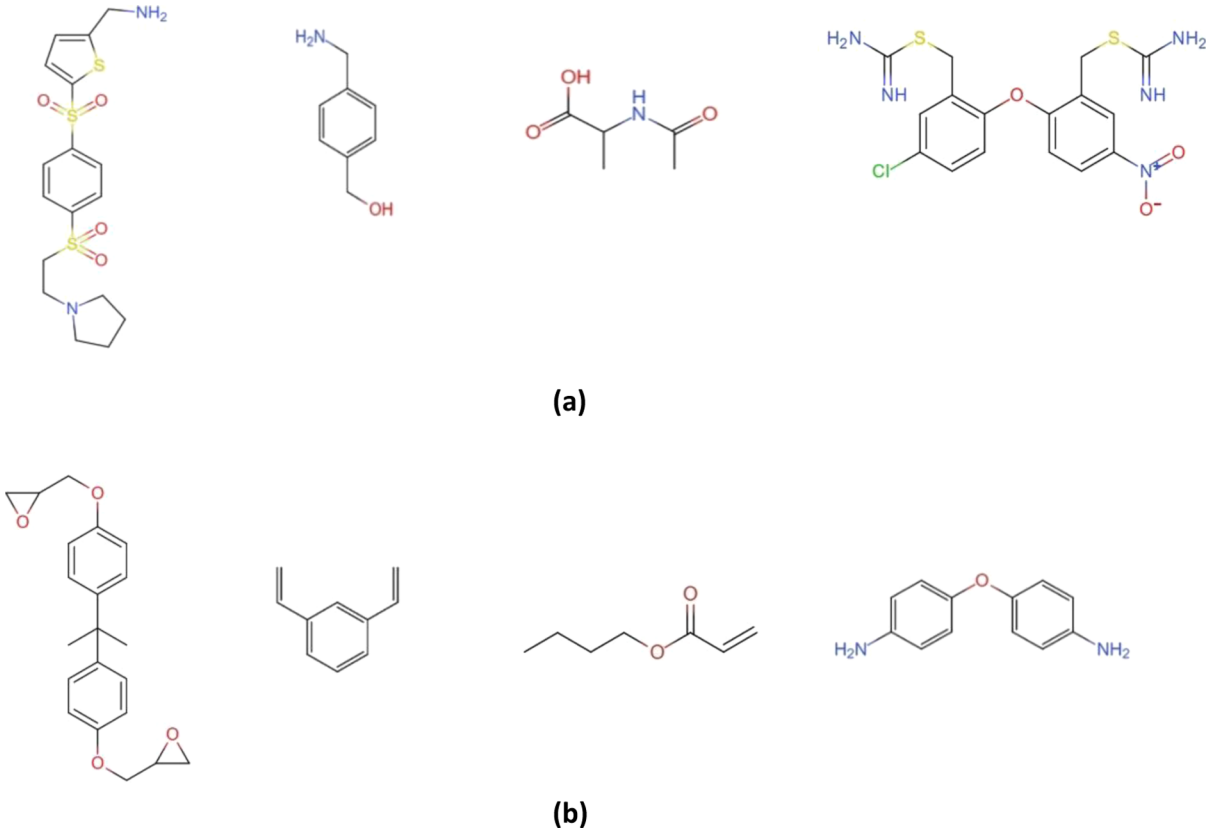


Figure 2. Comparison between (a) drug molecule and (b) TSMP monomer or cross-linker.

only thousands of SMP data can be collected from this database. If one further narrows the search to "TSMP", the search engine can only return about 200 papers at the time of this paper submission. As for "UV-curable shape memory polymer", only 22 articles can be returned. On the contrary, as indicated above, it is not difficult to find the public databases for drug molecules. From the aspect of chemistry, both the drug molecule and TSMP monomers belong to small molecules, which share many similarities in the chemical structures (see Figure 2). Therefore, we propose a new advanced ML approach "transfer learning−variational autoencoder (TL-VAE)" to tackle this problem. Specifically, we first let the ML model learn some common characteristics from a good amount of drug molecules, which allows us to create a pretrained model or a raw model. After that, by continuously learning the characteristics in a small TSMP data set, the model can be fine-tuned and is able to excellently capture the detailed features of TSMP monomers. Meanwhile, because more data are used in the learning process, overfitting can be alleviated to some extent. Additionally, on one hand, according to Lee et al.,[46] the premise to establish a good ML model is that the output loss (the difference between the prediction outcome and ground truth) should be Gaussian random distribution, which sets a high standard for the data size. On the other hand, it is worth noting that all the reported results are inevitably biased and do not meet the Gaussian random distribution. This is because although the experimental results are supposed to yield a Gaussian distribution, all research studies tend to report selected results, which suggests that only some particular SMPs were reported; for example, only the SMPs with large recovery stress were reported but the major part (the SMPs with moderate recovery stress) were involuntarily ignored. Therefore, through the introduction of a large number of small molecules in the database, which can form a Gaussian random distribution with a great probability, this issue can be solved in TL-VAE. Last but not the least, our approach also overcomes the intrinsic limitation of the VAE itself. To be specific, the VAE is more data-hungry than the general ML model. In order to make the VAE perform properly, at least thousands of data points are required. For example, Lee et al. collected 5473 alloy entries to establish a modified VAE;[46] Gómez-Bombarelli et al. adopted 250,000 drug-like molecules to train their VAE.[47] Apparently, the number of data points for the VAE has significantly exceeded the available TSMP data. Fortunately, the application of transfer learning enables the VAE to learn features from a large drug data set and hence put it into practice. In a nutshell, by applying TL-VAE, our approach solved the problems facing TSMP—the lack of a training data set, Gaussian distribution, and normal operation of the VAE.

Regarding the third difficulty, that is, the absence of the molar ratio, it has been a common limitation in the previous studies involving the ML for polymers. For example, Miccio and Schwartz adopted a CNN to predict T_g for polymers, but wherein only the polymers that composed of a single monomer were taken into consideration.[48] Wu et al. developed a ML framework that is able to predict the thermal conductivity of polymers by using Bayes' theorem, but wherein they only dealt with the homopolymer and circumvented the issue of the molar ratio.[49] In these previous studies, all the polymers were homopolymers, hence one monomer can represent the whole polymer network and there is no need to introduce the molar ratio. However, most TSMPs belong to copolymers; thus, the molar ratio becomes an unavoidable issue. Meanwhile, from

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521

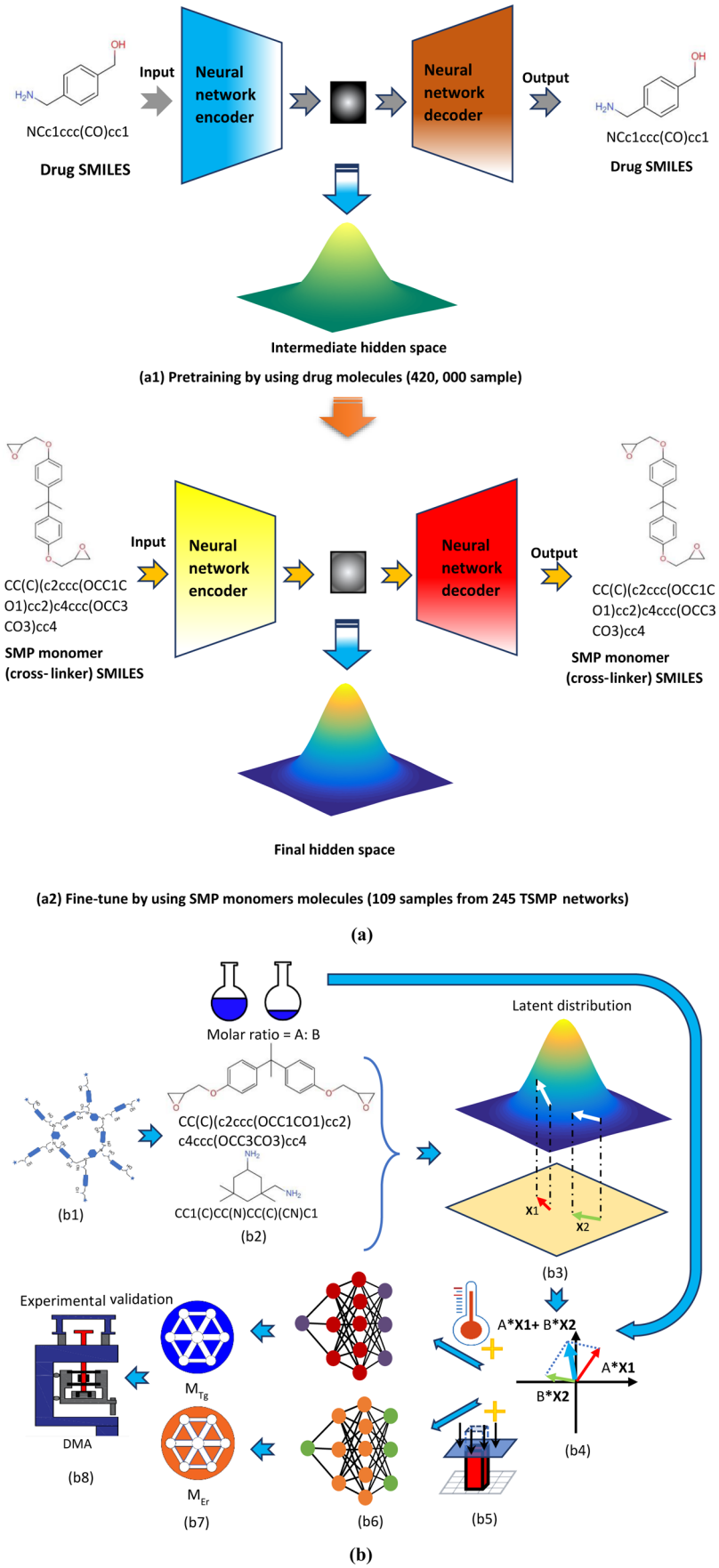


Figure 3. continued

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521

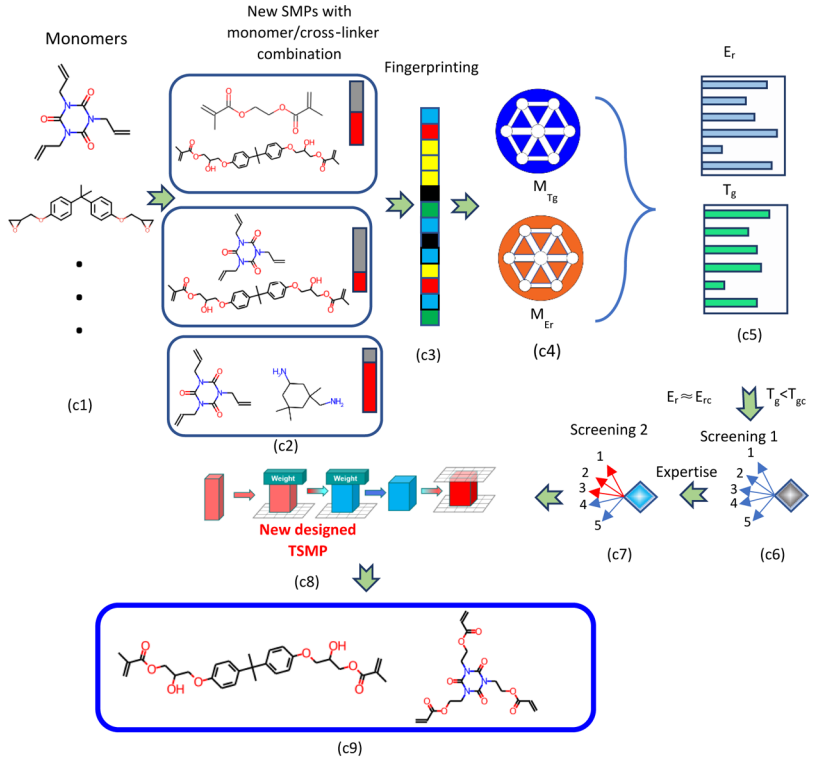


Figure 3. Pipeline of the new TSMP design. (a) TL-VAE model establishment process, (b) TSMP properties model establishment process (*A* and *B* represent the respective molar percentages in the TSMP network), and (c) new TSMP discovery.

the point of view of chemistry, for certain monomers and cross-linkers, a network with the highest cross-linking density can be formed if the molar ratio accords with the stoichiometric ratio. Moreover, if the molar ratio does not accord with the stoichiometric ratio but consecutively changes around the stoichiometric ratio, the cross-linking density would gradually increase or decrease, resulting in a continuously varying domain of the polymer properties,[42,50] which can be partly explained by the previous thermomechanical models. As suggested by the thermomechanical model of the Vogel−Fulcher−Tamman law based on entropy,[51] for a polymer chain with only one bond type and constant length, the stiffness of the polymer chain is inversely proportional to the number of polymer molecules. In other words, the greater the number of the SMP molecules present in a chain is, the softer the chain is. Hence, different molar ratios lead to different numbers of molecules in a polymer chain and further induce distinct moduli. These continuous property changes should be captured by the ML model.

In this study, we applied a combination of the VAE and weighted-vector combination method (WVCM) to capture the TSMP property changes that are caused by the molar ratio change between the monomer and cross-linker. Meanwhile, the VAE provides a probabilistic approach for describing the observation in a latent space; based on the VAE, we further developed a new fingerprinting method to encode the TSMP network, that is, the WVCM. It will be demonstrated later that the continuous property changes with the molar ratio are reasonably represented by this method. The advantage of this method is that it can fingerprint the TSMP network into a continuous vector, which can reasonably embody the continuous changes of the TSMP network with different molar ratios.

The aim of this study is to develop an enhanced ML framework that is able to discover new TSMPs or improve the performance for the known TSMPs. The paper is organized as follows. In Section 2, a VAE model is built by converting the TSMP monomer chemical structures into multiple-dimensional vectors in a two-step training process. Next, by combining the multiple-dimensional vectors and WVCM, prediction models for TSMP properties (T_g and rubbery modulus E_r) are established. We then validate them by experiments. Furthermore, in Section 3, five new types of TSMPs with desired properties are discovered with further experimental validation. Finally, in Section 4, some important conclusions are drawn.

2. METHODS

In this section, we present three main procedures in the framework, that is, the establishment of the two-step VAE model, establishment of the TSMP properties model, and establishment of the TSMP forward design model. The pipeline for discovering new TSMPs with the assistance of ML is shown in Figure 3. The detailed procedures are described as follows:

(1) Establishment of the TL-VAE model.
   (i) The VAE model is first trained by 420,000 drug molecules, and an intermediate hidden space is obtained in (a1).
   (ii) The VAE model is fine-tuned by 109 monomers, and then, the final hidden space is obtained in (a2).

(2) Establishment of the TSMP properties model.
   (i) According to the monomers (b2) that constitute a TSMP network (b1), the TSMP monomers can be

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521

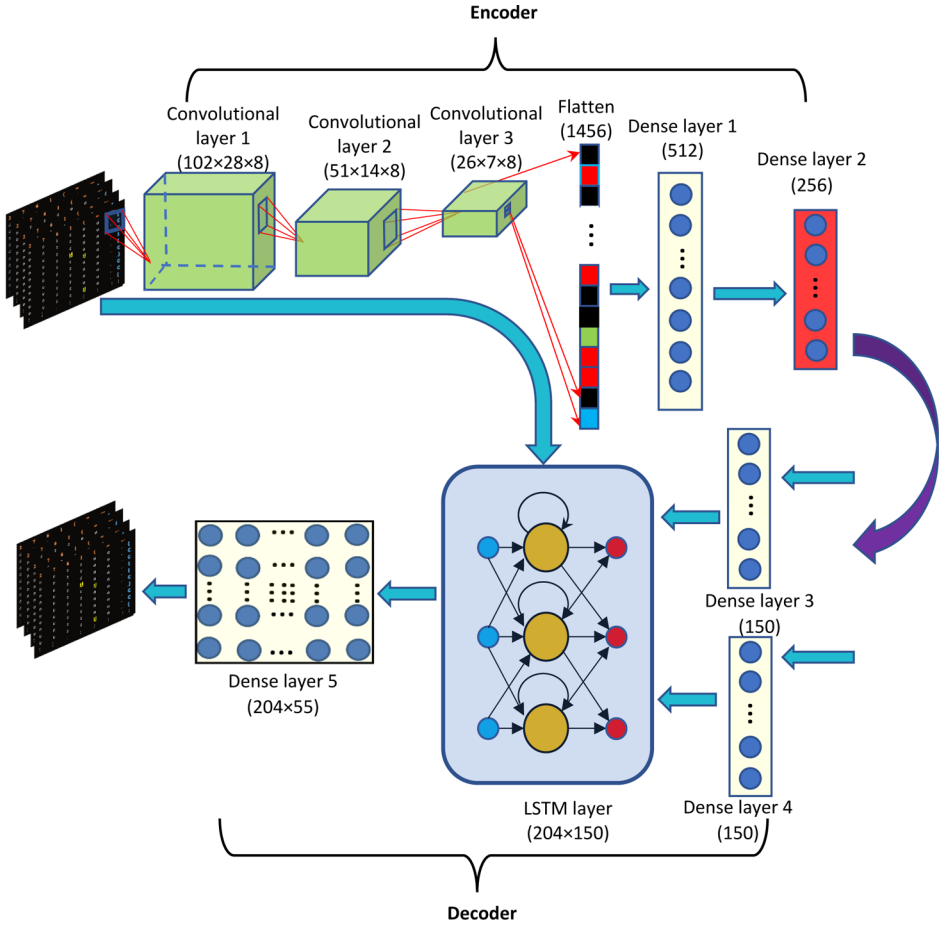


Figure 4. Basic pipeline structures for the network of a VAE model (the number in the bracket represents the number of neurons and the red color of the dense layer represents the output of the latent space).

fingerprinted into the corresponding high-dimension vectors in the hidden space by the VAE (b3).

(ii) According to the molar ratio, the high-dimension vectors for the monomers produce a weighted high-dimension vector for the TSMP network according to the WVCM in (b4).

(iii) Combining the high-dimension vectors and the corresponding properties data for the TSMP networks (T_g and E_r), two models can be further established in (b7), that is, the glass transition temperature model M_{T_g} and the rubbery modulus model M_{E_r}. The established models are validated by the experimental test (b8).

(3) Establishment of the TSMP forward design model.

(i) Collecting all monomers and cross-linkers in the training data set (c1), which are used to generate random arrangements and then produce new TSMP networks (c2).

(ii) The new possible TSMP networks (c2) can be fingerprinted into high-dimension vectors by the hidden space and different molar ratios.

(iii) The T_g and E_r for the new TSMP networks (c5) can be predicted by M_{T_g} and M_{E_r} in (c4).

(iv) Two screening procedures are conducted, that is, the samples (c5) can be screened by the benchmark (c6) and the chemical expertise (c7) in succession.

(v) The screened results are validated by experiments (c8), and finally, we can obtain the desired TSMPs (c9).

2.1. VAE Model. The VAE model is a type of recently developed autoencoder-like architecture, which provides a probabilistic underlying causal relation between the input and latent space. The causal relation has the potential for generalizability. It was first defined by Kingma and Welling in 2013[52] and since then has found wide applications in many research fields.[46,47,53,54] In the field of new material discovery, the VAE also came into use in some studies. For example, by combining the VAE and nondominated sorting genetic algorithm, Lee et al. pinpointed several novel thermomechanically controlled and processed steel alloys, and the predictions agree with the rule-based thermodynamic calculation tool.[46] Samanta et al. proposed a novel VAE for molecular graphs, wherein the encoder and decoder were optimized by several technical innovations, and the decoder was able to generate the molecules with the spatial coordinates of the atoms in it.[54]

The main advantage of the VAE is that it is able to learn a smooth latent state representation or vector through the input data and hence has a potential to form a generative model. As shown in Figure 3a, we first adopted canonicalized simplified molecular input line-entry system (SMILES) linear notions to represent the chemical structures, which were then transformed to the binary 204 × 55 matrices by a one-hot process (see the details in Figure 3 in ref 9.), wherein 204 and 55 represent the maximum length of the SMILES and the dictionary length, respectively. The generated binary matrices can be viewed as gray images. Since it can be treated an image, it is reasonable for us to leverage the CNN. This combination has been commonly adopted by other researchers.[48,55−57] Next, the matrices were input into the VAE model. Our VAE consists of two neural networks, that is, an encoder and a decoder. Among them, the encoder aims to convert the matrices to the representative vectors with 256 dimensions, and the decoder aims to restore the representative vectors into the original 204 × 55 matrices. It should be mentioned that the encoder and decoder can be the same or have different architectures; for example, an encoder composed of fully connected layers and a decoder composed of fully

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521

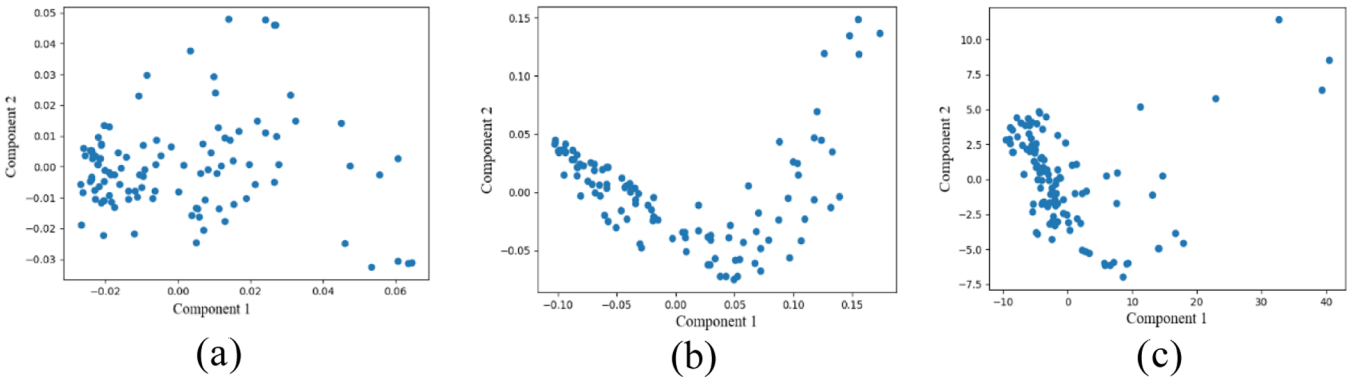


Figure 5. Three PCA mappings by adopting different training data sets. (a) 109 monomer (or cross-linker) SMILES; (b) 137,945 monomer SMILES (by exhaustively writing all possible 109 monomers); and (c) 420,000 drug molecules + 137,945 monomer SMILES (by exhaustively writing all possible 109 monomers).

connected layers,[46] an encoder consisting of a CNN layer and dense layers and another decoder consisting of a CNN layer and dense layers,[7] an encoder consisting of a CNN layer and dense layers and a decoder consisting of a recurrent neural network (RNN) layer and dense layers,[47] and so on. In this study, the VAE model comprises an encoder consisting of a CNN layer and dense layers and a decoder consisting of a long short-term memory (LSTM) neural network layer and dense layers. The basic pipeline structure of the network is shown in Figure 4. In this model, the CNN is able to learn the features of the images[58−60] and then capture the characteristics from the input matrices; the LSTM is able to learn the order dependence of the sequences[61−63] and then reconstruct the initial matrices. We choose the LSTM layer as the core part of the decoder. The reason is that it has a longer memory to the sequence than the common RNN and hence can learn the correct syntax for the long SMILES with higher possibility. It should be mentioned that we leveraged the "stateful" LSTM in the training, which can learn the relevance of letters in a SMILES.

As indicated above, the chemical structures of TSMPs are relatively complex, but we only collected 109 unique monomers (from 245 TSMPs), which hardly makes the VAE learn the syntax of the SMILES, not to mention the particular atomistic orders for the SMILES of the TSMP monomers. Therefore, we employed a two-step training approach. First, we collected 420,000 small molecules from an open database ChEMBL[64] (a biological database with small drug molecules). With the huge amount of data in it, the VAE can learn not only the syntaxes of SMILES and the basic chemical structures of small drug molecules but also the particular atomistic orders of the small drug molecules with similar chemical structures to TSMP monomers. Second, we fine-tuned the models and further enabled them to be familiar with TSMP monomers. Considering only 109 unique TSMP monomers were present in this small data set, we applied data augmentation. Specifically, an open-source cheminformatics software RDKit was used to exhaustively provide all possible SMILES linear notions for a concatenated SMILES, which expanded our TSMP monomers to 137,945 different linear notions. It should be mentioned that the different notions for the same chemical structure could only be slightly different. For example, a chemical structure with the concatenated SMILES "C=C(C)C(=O)-OCCOCCOCCOCCOCCOCCOCCOCCOC(=O)C(=C)C" can also be written as the nonconcatenated SMILES linear notions, including "C(COCCOCCOCCOCCOCCOCCOC(=O)C(C)=C)OCCOC(=O)C(=C)C", "C(OCCOCCOC(=O)C(C)=C)-COCCOCCOCCOCCOCCOCCOC(=O)C(=C)C", "C-(OCCOCCOCCOCCOCCOCCOCCOCCOCCOC(C(=C)C)=O)(=O)C(=C)C", and so forth. In addition, in the second-step training, we only trained the weights and biases for the last two layers and froze other layers, that is, the LSTM layer and the last dense layer. It should be noted that the VAE can be used for optimization. In this study, we did not utilize this function of the VAE. The detailed reasons can be found in Supporting Information 1. However, we show some possible VAE-generated new chemical structures in Supporting Information 2.

In order to prove the necessity of our two-step training, we mapped the chemical structures of the TSMP monomer to 2D vectors using the VAE and principal component analysis (PCA). In this study, the mapped high-dimension vector is important because the subsequent 1D CNN is based on the VAE results. We hope that the VAE model can map some really important features into the latent space, which will be beneficial for the later 1D CNN prediction. As indicated by Asperti,[65] the sparsity can "force the model to focus on the really important features, highly reducing the risk of overfitting"; thus, it is important to use the VAE. A similar scenario can be found in our study; see Figure 5. Specifically, we compared three PCA mappings by adopting different training data sets, that is, (1) 109 monomer SMILES, (2) 137,945 monomer SMILES (by exhaustively writing all possible 109 monomers), and (3) the combination of 420,000 drug molecules and 137,945 monomer SMILES (by exhaustively writing all possible 109 monomers). It can be found that a larger data set indeed can be mapped into a sparse space; thus, we believe that our strategy, similar to that of Asperti,[65] has captured more important features from monomers (or cross-linkers) and is successful. Furthermore, through the prediction comparison for the three training data sets, we found that our strategy is the best in all the three models (see Table S1 in Supporting Information 3).

The loss function used here is "categorical cross-entropy", which reads as follows

$$L(\hat{y}, y) = -\frac{1}{N}\sum_{i}^{N} [y_i \log \hat{y}_i + (1 - y_i)\log(1 - \hat{y}_i)] \quad (1)$$

where $\hat{y}_i$ and y_i are the prediction outcome and ground truth, respectively, and N is the number of the training data. The hyperparameters used in the VAE model are listed in Table 1. The parameter optimization is shown in Supporting Information 4.

Table 1. Hyperparameters Used for the VAE Model

hyperparameters names	value or item
batch size for step 1−2	1843, 2457
learning rate	0.005
number of filters 1−3	8, 8, 8
kernel size 1−3	3, 3, 3

Since our subsequent 1D CNN is based on the VAE model, the converted high-dimensional vector is vital to the prediction. As is well-known, if the VAE is well-trained, it is able to map the input into the same output, namely,

$$D(E(x)) = x \quad (2)$$

where

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521

$$E = q(z|x), \quad D = q(x|z) \tag{3}$$

in which E and D are the mappings of the encoder and decoder, respectively; x and z are the input and output; and q denotes mapping. Therefore, the similarity between the input SMILES and output SMILES should be considered as a key index to evaluate the performance of the VAE. Here, we introduce the average cosine similarity as

$$\mathrm{sim}_{\mathrm{avg}} = \frac{\sum_{i=1}^{n} \frac{\mathbf{v}_{\mathrm{inp}} \cdot \mathbf{v}_{\mathrm{out}}}{\| \mathbf{v}_{\mathrm{inp}} \| \| \mathbf{v}_{\mathrm{out}} \|}}{n} \tag{4}$$

where $\mathbf{v}_{\mathrm{inp}}$ and $\mathbf{v}_{\mathrm{out}}$ represent the binary matrices corresponding to the input SMILES and output SMILES, respectively.

It should be mentioned that all the matrices have been reshaped to the vectors. To validate that our choice for training data is reasonable, we employed three different training data sets for comparison. As given in Table 2, by adopting the two-step training, the average similarity can reach 92.39%, which is chosen as our training data set for the VAE. It proves the superiority of our strategy. After the training, we input 109 unique TSMP monomers to test the performance of the VAE model and found about 80% of outputs can fully accord with the SMILES syntax. It makes sense because part of the information is lost in the encoding process and cannot be completely restored in the decoding processes. For the remaining 20% TSMP monomers, although the VAE cannot fully output the perfect SMILES linear notions, some important features of the chemical structure still remain. In order to justify it, we compared two initial SMILES to the decoder outputs with the wrong syntax in Table 3. It can be seen that if we make a slight modification for the decoder output (only one letter is changed), the linear notation can accord with the standard SMILES syntax. Upon taking a closer look at the chemical structures represented by the modified SMILES syntax, we find that some features of the original chemical structure still hold.

Table 2. Average Cosine Similarity for Three Different Training Data Sets

training data set	average similarity (%)
109 monomer (or cross-linker) SMILES	8.80
137,945 monomer SMILES (by exhaustively writing all possible 109 monomers)	76.84
420,000 drug molecules + 137,945 monomer SMILES (by exhaustively writing all possible 109 monomers)	92.39

2.2. TSMP Properties Model. Because the VAE model is established, we can further obtain the fingerprints of the TSMP network. The method is created by us and is named as the "weighted-vector combination method". It can be divided into two procedures:

(i) Calculate the representative vectors $\{\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, ...\mathbf{m}_n\}$ for the TSMP monomers in a TSMP network via the VAE.

(ii) Obtain a new high-dimensional resultant vector by combining the monomers' respective molar ratio and vectors, that is,

$$\mathbf{S} = \mathbf{m}_1 \cdot a_1 + \mathbf{m}_2 \cdot a_2 + \mathbf{m}_3 \cdot a_3 + ...\mathbf{m}_n \cdot a_n \tag{5}$$

where a_i $(i = 1, 2, ...n)$ represents the molar percentage in the whole TSMP network. Based on the WVCM, every TSMP network can be converted into a vector with 256 dimensions. It accords with the chemical mechanism, that is, no matter how many monomers and cross-linkers participate in the reaction, only one SMP network can be finally produced.

Based on this, the TSMP properties prediction models can be further built for 245 SMP data. To increase the robustness, two ML methods are used here, that is, CNN learning and support vector machine (SVM) learning.

2.2.1. CNN Learning. As shown in Figure 6, we adopted multiple convolution 1D layers to learn two mappings, that is, Ψ: $\mathbf{S} \rightarrow T_g$, the mapping between the representative vector $\mathbf{S}$ and T_g, and Φ: $\mathbf{S} \rightarrow E_r$, the mapping between the representative vector $\mathbf{S}$ and rubbery modulus E_r. The hyperparameters for the TSMP properties model are given in Table 4. The parameter optimization is shown in Supporting Information 4.

The loss function used is the "MAPE", which can be written as

$$\mathrm{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{6}$$

where $\hat{y}_i$ and y_i are prediction outcome and ground truth, respectively, and n is the sample number of the training data.

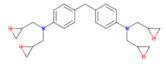
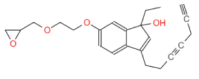
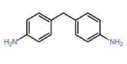
2.2.2. SVM Learning. SVM is a set of supervised ML method, which is one of the most robust prediction methods based on statistical leaning. The goal of SVM is to obtain the hyperplane that satisfies the condition

$$\mathbf{w}^{\mathrm{T}}\mathbf{x} - b = 0 \tag{7}$$

where $\mathbf{w}$ is the normal vector to the hyperplane, $\mathbf{x}$ is a p-dimensional vector, and b is a bias term. The aim of SVM is to minimize

$$\left[\frac{1}{n} \sum_{i=1}^{n} \max(0, 1 - y_i(\mathbf{w}^{\mathrm{T}}\mathbf{x}_i - b)) \right] + \lambda \|\mathbf{w}\|^2 \tag{8}$$

Table 3. Comparison between Original SMILES and the Decoder Output with the Wrong Syntax

No.	Original SMILES and chemical structures		Decoder output with wrong syntax	Slightly modified decoder output
1	SMILES	c3cc(N(CC1CO1)CC2CO2)ccc3Cc6ccc(N(CC4CO4)CC5CO5)cc6	c1c(CC)(O)c2cc(OCCOCC3CO3)ccc2c1CCC#C)CC#C	c1c(CC)(O)c2cc(OCCOCC3CO3)ccc2c1CCC#CCC#C
	Chemical structure		NA	
2	SMILES	Nc2ccc(Cc1ccc(N)cc1)cc2	Nc2ccc(C)cc2)c1c(C)cccc1	Nc2ccc(C)cc2Cc1c(C)cccc1
	Chemical structure		NA	

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521

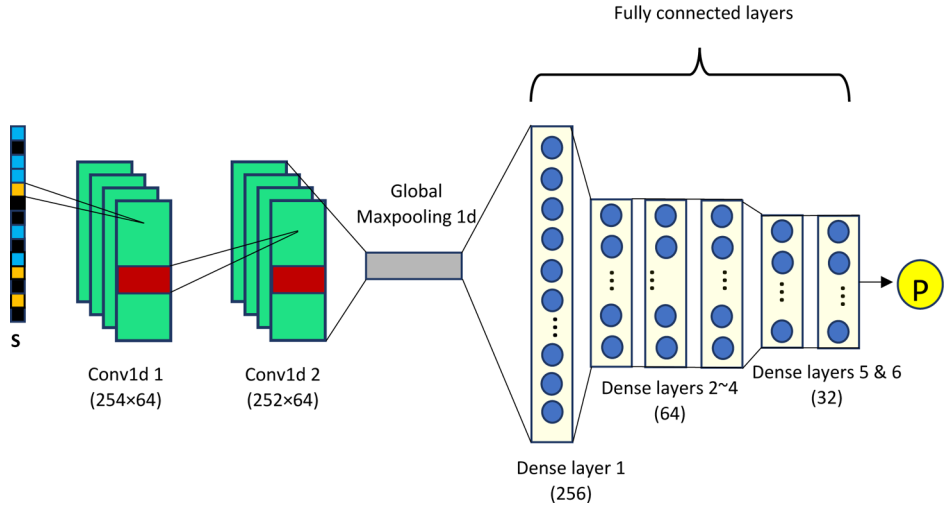


Figure 6. Basic structure for the neural network of the TSMP properties prediction model.

Table 4. Hyperparameters Adopted in the Training of the TSMP Properties Model

hyperparameter name	value or item
ratio between the training data and test data	80/20 (random_state=7)
batch size	256
learning rate	0.01
number of filters 1−2	64, 64
kernel size 1−2	3, 3
neuron number in hidden layers	256, 64, 64, 64, 32, 32

where λ is a parameter that determines the trade-off between increasing the margin size and ensuring that x_i is located on the right side of the margin and n is the number of vectors.

Besides the MAPE, we also introduced the "percentage of correct point" (PCP) as another evaluation criterion. It should be noted that "correct" means that the difference between the prediction and the ground truth is within a reasonable margin. Because the order of magnitude covered by T_g and E_r are totally different, that is, the T_g of TSMP ranges from 20 to 280 °C and only 1 order of magnitude is covered, and the rubbery modulus of TSMP varies from 0.38 to 280 MPa and almost covers 3 orders of magnitude, we applied two distinct criteria, that is, the percentage error for T_g within 15% is considered as "correct" and we let the percentage error for E_r within 30% to be considered as "correct".

In order to demonstrate the superiority of our strategy, we introduced a benchmark for comparison. Specifically, we directly leveraged label encoding to generate vectors for monomers and cross-linkers (and the others follow the same methods) and then adopted SVM for prediction. It can be seen in Table 5 that our new model exhibits a lower MAPE and a higher PCP in all the eight comparisons for SVM prediction. Meanwhile, we also list the results of our previous model as an indirect comparison. Although the databases for these two studies are not completely the same, the corresponding linear notations (one is SMILES and the other is BigSMILES) and predicted targets are similar. At the same time, we provided a circular topological fingerprint as another baseline model, that is, Morgan fingerprinting or extended-connectivity fingerprinting.[66] In this fingerprinting, all the molecular structures of monomers or cross-linkers are converted into binary vectors according to their structures, and then, TSMPs can be converted into vectors with the same dimension through the WVCM. Here, without losing generality, we adopt the circular radius of $R = 2$ and three different dimensions, that

Table 5. Comparison of the Prediction Discrepancies for the VAE + CNN, VAE + SVM, Label Encoding + SVM, Morgan Encoding + SVM, and the CNN Model in Our Previous Study[9] (Here, Label Encoding Uses the Same Dictionary as That of Our New Model, and *B* and *R* Represent the Dimension of the Vector and Circular Radius in Morgan Encoding, Respectively)

types of models	ML method	model output	MAPE in training data (%)	PCP in training (%)	MAPE in test data (%)	PCP in test (%)
our new model	VAE + CNN	M_{Tg}	7.38	91.33	13.91	71.43
		M_{Er}	15.64	89.80	27.15	71.43
	VAE + SVM	M_{Tg}	19.91	86.74	17.50	75.51
		M_{Er}	34.21	71.94	36.42	44.90
direct label encoding	vector encoding + SVM	M_{Tg}	31.87	64.29	24.82	63.26
		M_{Er}	50.01	55.10	52.09	30.61
Morgan encoding (B = 256, R = 2)	Morgan encoding + SVM	M_{Tg}	20.86	73.46	21.76	65.30
		M_{Er}	34.02	67.85	61.06	42.86
Morgan encoding (B = 1024, R = 2)	Morgan encoding + SVM	M_{Tg}	20.08	73.47	21.15	69.38
		M_{Er}	33.11	68.88	62.39	44.90
Morgan encoding (B = 2048, R = 2)	Morgan encoding + SVM	M_{Tg}	19.60	73.98	19.83	69.39
		M_{Er}	33.08	68.37	60.05	40.82
Morgan encoding (B = 2048, R = 4)	Morgan encoding + SVM	M_{Tg}	18.45	75.51	19.34	73.47
		M_{Er}	31.49	69.90	56.16	44.898
old model[9]	CNN	M_{Tg}	3.12		28.33	
		$M_{\sigma s}$	9.81		27.41	

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521

Table 6. Comparison between Experimental Results and ML Model Predictions

epoxy network number	molar ratio (EPON862:4-APD:PEI)	approach to obtain data	T_g (°C)	discrepancy (%)	E_r (MPa)	discrepancy (%)
1	0.64:0.04:0.32	experiment	159.85	17.32	14.54	30.66
		ML model	187.54		19.00	
2	0.69:0.07:0.24	experiment	151.49	23.77	21.25	20.00
		ML model	187.50		16.99	
3	0.74:0.10:0.16	experiment	147.78	19.02	21.85	36.45
		ML model	175.89		13.89	
4	0.83:0.17:0	experiment	135.7	10.56	36.57	11.41
		ML model	121.37		32.40	
average discrepancy value	17.68%	24.64%				

is, B = 256, 1024, and 2048. Meanwhile, considering that TSMP monomers could possess more complex structures than common drug molecules, we also considered a larger circular radius, that is, R = 4. From Table 5, it is found that the performance of the model based on Morgan fingerprinting improves slightly upon increasing the vector dimension B and circular radius R. However, even for the model based on Morgan fingerprinting with the best performance (B = 2048 and R = 4), our new model still takes the lead in six out of the eight comparisons. As for the other two comparisons, our new model is just slightly behind. More importantly, by comparing the error differences for rubbery modulus between the training data and test data of the models based on Morgan fingerprinting, we found that all the differences exceed 20%, which indicates significant overfittings (while it does not happen in our new model). Therefore, the direct combination of Morgan fingerprinting and the molar ratio could be inappropriate for the small TSMP data set in this study. As for VAE + SVM and VAE + CNN models, it is clearly shown that the two models have similar prediction accuracies for T_g and E_r but the VAE + CNN model shows better performance for some parameters; hence, we used VAE + CNN to design our new TSMP in the following section. More importantly, the VAE + CNN model in this study shows significant improvement over our previous model. Specifically, our new model reduces the MAPE by more than half (13.91 vs 28.33%) for M_{Tg}; thus, the accuracy is enhanced. Meanwhile, for a bigger data set (the data size increases by 2 times), this new model has almost the same accuracy as that of the recovery stress model for M_{Er} or $M_{\sigma s}$ (27.15 vs 27.41%); thus, the new VAE + CNN model is robust. In addition, it can be observed that the MAPE differences between the training data and test data are only 6.53 and 11.51%, respectively. Both of them are much less than 20%; thus, the overfitting problem has been significantly alleviated. Furthermore, our VAE + CNN model prediction is better than the benchmark prediction using the direct encoding + SVM model, further validating the advantage of using VAE + CNN in this study.

With the mapping Φ: $\mathbf{S} \rightarrow E_r$, we can further estimate the maximum recovery stress for a new TSMP network under a small deformation. According to Wang and Li,[38] the recovery stress can be understood as the combination of four components, that is, residual stress σ_{res}, memorized stress σ_{mem}, thermal stress σ_T, and relaxed stress σ_{rel}

$$\sigma_r = \sigma_{res} + \sigma_{mem} + \sigma_T - \sigma_{rel} \quad (9)$$

Besides, as indicated by Yan et al.,[9] if the programming strain ε_{prog} is small (≤15%), the residual stress σ_{res} and relaxed stress σ_{rel} can be omitted, and the memorized stress can be written as

$$\sigma_{mem} = E_r\varepsilon_{prog} \quad (10)$$

Combining eq 9 with eq 10, the recovery stress can be estimated as

$$\sigma_r = (E_r\varepsilon_{prog})\cdot R_{fix}\cdot R_{re} \quad (11)$$

where R_{fix} and R_{re} represent the shape fixity ratio and shape recovery ratio, respectively. If the TSMP possesses a good SME, that is, $R_{fix} \approx R_{re} \approx 1$, the recovery stress can be approximately calculated as the product between the rubbery modulus E_r and programming strain ε_{prog}:

$$\sigma_r = E_r\varepsilon_{prog} \quad (12)$$

In other words, a larger rubbery modulus usually means larger recovery stress. Additionally, another advantage for this method is that more data can be collected. This is because only a few studies provide the information for recovery stress, while almost all of them report the rubbery modulus value.

In addition, in order to realize the application in 3D/4D printing, it is popular to immerse the SMP matrix into a solvent before printing, which may significantly influence the T_g.[67,68] Therefore, to improve this study, we plan to incorporate the solvent effect in our future studies.

2.3. Experimental Validation. In order to further validate the new ML model, we synthesized four types of new epoxy networks (Table 6) with the same three monomers (EPON862, 4-APD, and PEI, see Figure 7) but different molar ratios between the monomers.

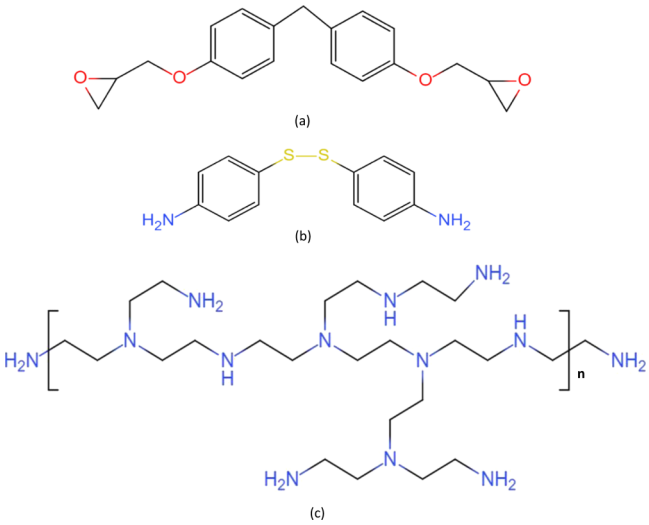
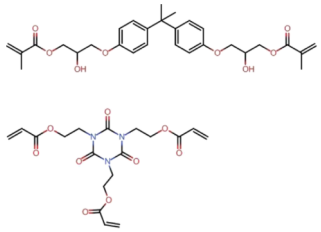
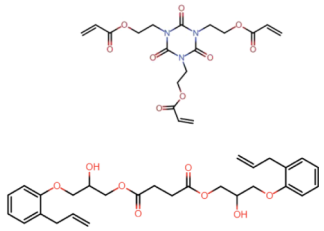
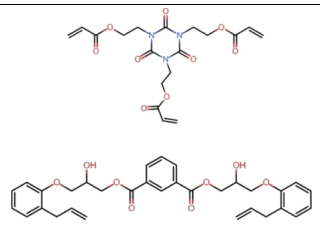
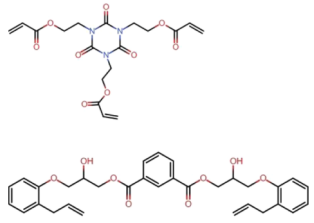
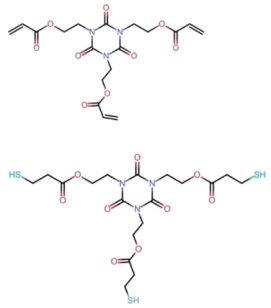


Figure 7. Chemical structures of the three monomers used in the synthesis: (a) EPON862, (b) 4-APD, and (c) PEI.

The synthetization process is similar to that in our previous work.[9] The T_g and E_r for the four types of TSMPs were measured by a dynamic mechanical analyzer (DMA), and the experimental results are compared with the ML prediction in Table 6. As indicated in Table 6, both M_{Tg} and M_{Er} show low MAPE values (the MAPE values for M_{Tg} and M_{Er} are 12.78 and 15.79%, respectively), which is basically in accordance with the errors for the test data and hence partially validates our model. Of these two properties, M_{Tg} shows better prediction again and captures a trend of T_g. That is, with the increase in the molar ratio between EPON 862 and 4-APE, the mobility of the polymer chain gradually reduces, and T_g increases, which is reasonably captured by the model. As for the rubbery modulus, it increases with the molar ratio when it is close to the ideal stoichiometric ratio and vice versa. To better learn this strategy, we

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521

Table 7. Chemical Structures, Predicted Glass Transition Temperatures T_g, and Rubbery Modulus E_r for Five New Promising TSMPs

NO.	Combination	Chemical structures of monomers	E_r (MPa)	T_g (°C)
1	BIS-GMA:TAI=0.1:0.9		344.24	246.74
2	TAI: AEG1=0.9:0.1		337.44	246.12
3	TAI: AEG2=0.9:0.1		351.77	251.32
4	TAI: AEG2=0.8:0.2		340.35	247.65
5	TAI: TMICN=0.7:0.3		319.66	245.98

expect to introduce some new features to reduce the prediction error in our future studies.

3. NEW TSMP DISCOVERY

By leveraging the newly developed ML model, it is possible for us to discover the desired TSMPs. In our database, the TSMP with the highest rubbery modulus was synthesized by Feng and Li,[29] with a value of 350 MPa. However, the limitation is that its T_g is too high, that is, T_g = 280 °C, thus limiting the application when the available trigger temperature is lower. For example, when the SMP is used for geothermal drilling,[69] it is expected that the recovery stress is high, but the trigger temperature is limited to about 220 °C so that the smart loss circulation materials can be driven by the geotherm. Therefore, we target to screen a new TSMP with a high rubbery modulus but lower T_g. It is noted that this is not a trivial task because usually, TSMPs with high recovery stress also have high T_g.

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521

Table 8. Comparison of Experimental Results between the New SMP and Previous SMPs with High Recovery Stress

references	glass transition temperature (T_g) (°C)	programming temperature (T_{pg}) (°C)	programming strain (ε_{pg}) (%)	stress recovery temperature (T_{rec}) (°C)	recovery stress (σ_{rec}) (MPa)	shape recovery ratio (R_f) (%)	curing type
new SMP	220.1	260	25	260	20.6	62	UV-curing
Li et al.'s SMP[32]	150	150	20	150	13.4	100	UV-curing
EPON-IPD[37]	141	150	45	170	17	89.6	thermal curing

We collected 109 TSMP monomers in the data set, which include 27 C═C monomers, 8 hydroxyl monomers, 7 carboxy cross-linkers, 5 thiol cross-linkers, 6 cyanate monomers, 21 epoxy monomers, 29 imine cross-linkers, and 6 other monomers (see Tables S5−S8 in Supporting Information 5). Among them, four combinations involving different monomers are chosen: (1) the combination of two C═C monomers, (2) the combination of one C═C monomer and one hydroxy cross-linker, (3) the combinations of one C═C monomer and one carboxy cross-linker, and (4) the combination of one C═C monomer and one thiol cross-linker. With nine molar ratios, that is, 0.1:0.9, 0.2:0.8, 0.3:0.7, 0.4:0.6, 0.5:0.5, 0.6:0.4, 0.7:0.3, 0.8:0.2, and 0.9:0.1, for any of the two-component TSMPs, the size of a new compositional space can reach 8019. Additionally, considering the prediction discrepancies of the ML models, we set the benchmarks as $T_g < 252$ °C ($\frac{T_g - T_{g_{new}}}{T_g} < 10\%$) and $E_r \approx 350$ MPa ($(|E_r - E_{r_{new}}|/E_r) < 9\%$). Finally, five types of new TSMPs were discovered, and the results are summarized in Table 7.

In order to further validate the ML model, we synthesized the No. 1 TSMP in Table 7 as described below. 10 g of tris[2-(acryloyloxy)ethyl]isocyanurate was heated at 80 °C until it was melted and was then mixed with 1.35 g of bisphenol A glycerolate dimethacrylate and 0.35 g of photoinitiator 2-hydroxy-2-methylpropiophenone by stirring for 20 min. The obtained mixture was degassed at 80 °C in a vacuum oven. The uniform mixture was filled into a polytetrafluoroethylene spacer with a thickness of 1.1 mm and clamped by two transparent plastic slides. Each side of the sample was cured in a UV chamber (IntelliRay 600, Uvitron International, USA) for 30 s under 50% irradiation intensity (232 nm, ~45 mW/cm^2). Its E_r and T_g values were measured by using the DMA, which were determined to be 383.60 MPa and 220.10 °C, respectively. The discrepancies between the experimental results and predictions are 10.26% and 12.10% for E_r and T_g, respectively. By comparing the new polymer and the TSMP by Feng and Li,[29] the T_g is reduced by 60 °C, while the high E_r is still maintained; thus, our initial goal of discovering new TSMPs with higher E_r and lower T_g has been achieved.

Next, in order to test the recovery stress for the new TSMP, a fully constrained stress recovery experiment was carried out. The experimental results are given in Table 8 and compared with those of the SMP with the highest recovery stress that has been discovered before. As shown in Table 8, when the programming temperature of the purely UV-curable TSMP is above T_g, the new TSMP is able to provide recovery stress much higher than the previous record[32] by 57%. It is noted that Li et al.[32] programmed and recovered the TSMP at the glass transition zone, while our new TSMP was programmed and recovered in the rubbery state, and thus, our new TSMP should have even higher recovery stress if we choose programming and recovery in the glass transition zone. Also, the recovery stress of our new TSMP exceeds that of another thermocured epoxy TSMP (EPON-IPD[37]) with high recovery stress by 21%. More recovery stress comparisons for UV-curable SMPs can be found in Supporting Information 6. Thus, this newly designed TSMP is consistent with our expectations for recovery stress and T_g, and our final design target is achieved. Additionally and notably, all the predicted molecules are reasonable and thermodynamically stable. This indicates that our ML model has learned some basic chemical principles about the covalence bond theory, such as the molecular orbital theory, Pauli's exclusion principle, and Hund's rules from the chemical language. On the basis of this knowledge, the ML model can voluntarily remove these thermodynamically unstable molecular structures and thus exhibit good accuracy and high efficiency of the prediction.

4. CONCLUSIONS

In summary, benefited from the huge database of drug molecules available, we have presented an enhanced ML approach based on a VAE model, two-step training, and the WVCM. Using this approach, we partially solved the two common existing problems for discovering TSMPs with ML, that is, lacking molar ratio information and limited training data. Comparing with our previous work, this approach has taken a major stride in both accuracy and robustness. By using this model, we screened five types of new UV-curable TSMPs and then validated one of them through synthesis and characterization. The comparison between the ML model and experimental results also shows good agreement. Therefore, the approach provides a promising framework to design and optimize new TSMPs. To our knowledge, the approach is the state-of-the-art in the TSMP field. It is believed that this approach is appliable not only to discover new TSMPs but also to discover other polymeric or nonpolymeric materials.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsami.1c20947.

Possible optimization by using the VAE; generation of new monomer (or cross-linker)-like chemical structures by adding Gaussian noise; prediction results from different training data sets; model parameter optimization; SMILES of monomers and cross-linkers; and comparison of recovery stress for UV-curable SMPs (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Guoqiang Li − *Department of Mechanical & Industrial Engineering, Louisiana State University, Baton Rouge, Louisiana 70803, United States;* orcid.org/0000-0002-

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521

7004-6659; Phone: 001-225-578-5302; Email: lguoqi1@lsu.edu

Authors

Cheng Yan − *Department of Mechanical & Industrial Engineering, Louisiana State University, Baton Rouge, Louisiana 70803, United States;* orcid.org/0000-0002-7567-0522

Xiaming Feng − *Department of Mechanical & Industrial Engineering, Louisiana State University, Baton Rouge, Louisiana 70803, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsami.1c20947

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the US National Science Foundation under grant number OIA-1946231, the Louisiana Board of Regents for the Louisiana Materials Design Alliance (LAMDA), the National Science Foundation under grant number 1736136, and the NASA cooperative agreement NNX16AQ93A under contract number NASA/LEQSF(2016−19)-Phase3-10.

■ REFERENCES

(1) Lopes, N.; Ribeiro, B. GPUMLib: An Efficient Open-Source GPU Machine Learning Library. *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.* **2011**, *3*, 355−362.
(2) Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-Based Learning Applied to Document Recognition. *Proc. IEEE* **1998**, *86*, 2278−2324.
(3) Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735−1780.
(4) Yeh, C.; Matsuda, N.; Huang, X.; Li, F. Temporal Convolutional Networks: A Unified Approach to Action Segmentation. *European Conference on Computer Vision*; Springer: Cham, 2016; pp 47−54.
(5) Guo, K.; Yang, Z.; Yu, C.-H.; Buehler, M. J. Artificial Intelligence and Machine Learning in Design of Mechanical Materials. *Mater. Horiz.* **2021**, *8*, 1153−1172.
(6) Mortazavi, B.; Podryabinkin, E. V.; Roche, S.; Rabczuk, T.; Zhuang, X.; Shapeev, A. V. Machine-Learning Interatomic Potentials Enable First-Principles Multiscale Modeling of Lattice Thermal Conductivity in Graphene/Borophene Heterostructures. *Mater. Horiz.* **2020**, *7*, 2359−2367.
(7) Kim, J.-Y.; Cho, S.-B. Deep CNN Transferred from VAE and GAN for Classifying Irritating Noise in Automobile. *Neurocomputing* **2021**, *452*, 395.
(8) Yan, C.; Li, G. Machine Learning Framework for Polymer Discovery. *Encyclopedia of Materials: Plastics and Polymers*; Elsevier, 2021.
(9) Yan, C.; Feng, X.; Wick, C.; Peters, A.; Li, G. Machine Learning Assisted Discovery of New Thermoset Shape Memory Polymers Based on a Small Training Dataset. *Polymer* **2021**, *214*, 123351.
(10) Shan, Q.; Song, J.; Zou, Y.; Li, J.; Xu, L.; Xue, J.; Dong, Y.; Han, B.; Chen, J.; Zeng, H. High Performance Metal Halide Perovskite Light-Emitting Diode: From Material Design to Device Optimization. *Small* **2017**, *13*, 1701770.
(11) Bradshaw, D.; Claridge, J. B.; Cussen, E. J.; Prior, T. J.; Rosseinsky, M. J. Design, Chirality, and Flexibility in Nanoporous Molecule-Based Materials. *Acc. Chem. Res.* **2005**, *38*, 273−282.
(12) Yount, W. C.; Loveless, D. M.; Craig, S. L. Small-Molecule Dynamics and Mechanisms Underlying the Macroscopic Mechanical Properties of Coordinatively Cross-Linked Polymer Networks. *J. Am. Chem. Soc.* **2005**, *127*, 14488−14496.
(13) Adcock, S. A.; McCammon, J. A. Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev.* **2006**, *106*, 1589−1615.
(14) Hegde, G.; Bowen, R. C. Machine-Learned Approximations to Density Functional Theory Hamiltonians. *Sci. Rep.* **2017**, *7*, 42669.
(15) Li, G.; Xu, W. Thermomechanical Behavior of Thermoset Shape Memory Polymer Programmed by Cold-Compression: Testing and Constitutive Modeling. *J. Mech. Phys. Solids* **2011**, *59*, 1231−1250.
(16) Yan, C.; Li, G. Design Oriented Constitutive Modeling of Amorphous Shape Memory Polymers and Its Application to Multiple Length Scale Lattice Structures. *Smart Mater. Struct.* **2019**, *28*, 095030.
(17) Yan, C.; Li, G. A Mechanism-Based Four-Chain Constitutive Model for Enthalpy-Driven Thermoset Shape Memory Polymers with Finite Deformation. *J. Appl. Mech.* **2020**, *87*, 061007.
(18) Yan, C.; Yang, Q.; Li, G. A Phenomenological Constitutive Model for Semicrystalline Two-Way Shape Memory Polymers. *Int. J. Mech. Sci.* **2020**, *177*, 105552.
(19) Shojaei, A.; Li, G. Thermomechanical Constitutive Modelling of Shape Memory Polymer Including Continuum Functional and Mechanical Damage Effects. *Proc. R. Soc. A* **2014**, *470*, 20140199.
(20) Wang, R.; Fang, X.; Lu, Y.; Yang, C.-Y.; Wang, S. The PDBbind Database: Methodologies and Updates. *J. Med. Chem.* **2005**, *48*, 4111−4119.
(21) Kim, S.; Thiessen, P. A.; Bolton, E. E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B. A.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. H. PubChem Substance and Compound Databases. *Nucleic Acids Res.* **2016**, *44*, D1202−D1213.
(22) Consortium, T. U. UniProt: the universal protein knowledgebase. *Nucleic Acids Res.* **2017**, *45*, D158−D169.
(23) Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Feng, Z.; Gilliland, G. L.; Iype, L.; Jain, S.; Fagan, P.; Marvin, J.; Padilla, D.; Ravichandran, V.; Schneider, B.; Thanki, N.; Weissig, H.; Westbrook, J. D.; Zardecki, C. The Protein Data Bank. *Nucleic Acids Res.* **2002**, *58*, 899−907.
(24) Butkiewicz, M.; Lowe, E.; Mueller, R.; Mendenhall, J.; Teixeira, P.; Weaver, C.; Meiler, J. Benchmarking Ligand-Based Virtual High-Throughput Screening with the Pubchem Database. *Molecules* **2013**, *18*, 735−756.
(25) Zhang, H.; Liao, L.; Saravanan, K. M.; Yin, P.; Wei, Y. DeepBindRG: A Deep Learning Based Method for Estimating Effective Protein-Ligand Affinity. *PeerJ* **2019**, *7*, No. e7362.
(26) Ong, E.; Wong, M. U.; Huffman, A.; He, Y. COVID-19 Coronavirus Vaccine Design Using Reverse Vaccinology and Machine Learning. *Front. Immunol.* **2020**, *11*, 1581.
(27) Belmonte, A.; Lama, G. C.; Gentile, G.; Cerruti, P.; Ambrogi, V.; Fernández-Francos, X.; De la Flor, S. Thermally-Triggered Free-Standing Shape-Memory Actuators. *Eur. Polym. J.* **2017**, *97*, 241−252.
(28) Santos, L.; Dahi Taleghani, A.; Li, G. Expandable Proppants to Moderate Production Drop in Hydraulically Fractured Wells. *J. Nat. Gas Sci. Eng.* **2018**, *55*, 182−190.
(29) Feng, X.; Li, G. High-Temperature Shape Memory Photopolymer with Intrinsic Flame Retardancy and Record-High Recovery Stress. *Appl. Mater. Today* **2021**, *23*, 101056.
(30) Li, G.; Xu, T. Thermomechanical Characterization of Shape Memory Polymer-Based Self-Healing Syntactic Foam Sealant for Expansion Joints. *J. Transp. Eng.* **2011**, *137*, 805−814.
(31) Feng, X.; Li, G. Versatile Phosphate Diester-Based Flame Retardant Vitrimers via Catalyst-Free Mixed Transesterification. *ACS Appl. Mater. Interfaces* **2020**, *12*, 57486−57496.
(32) Li, A.; Fan, J.; Li, G. Recyclable thermoset shape memory polymers with high stress and energy output via facile UV-curing. *J. Mater. Chem. A* **2018**, *6*, 11479−11487.
(33) Meng, H.; Li, G. A Review of Stimuli-Responsive Shape Memory Polymer Composites. *Polymer* **2013**, *54*, 2199−2221.
(34) Li, A.; Challapalli, A.; Li, G. 4D Printing of Recyclable Lightweight Architectures Using High Recovery Stress Shape Memory Polymer. *Sci. Rep.* **2019**, *9*, 7621.

(35) Lan, X.; Liu, Y.; Lv, H.; Wang, X.; Leng, J.; Du, S. Fiber Reinforced Shape-Memory Polymer Composite and Its Application in a Deployable Hinge. *Smart Mater. Struct.* **2009**, *18*, 024002.
(36) Scalet, G. Two-Way and Multiple-Way Shape Memory Polymers for Soft Robotics: An Overview. *Actuators* **2020**, *9*, 10.
(37) Fan, J.; Li, G. High Enthalpy Storage Thermoset Network with Giant Stress and Energy Output in Rubbery State. *Nat. Commun.* **2018**, *9*, 642.
(38) Wang, A.; Li, G. Stress Memory of a Thermoset Shape Memory Polymer. *J. Appl. Polym. Sci.* **2015**, *132*, 42112.
(39) Castro, F.; Westbrook, K. K.; Long, K. N.; Shandas, R.; Qi, H. J. Effects of Thermal Rates on the Thermomechanical Behaviors of Amorphous Shape Memory Polymers. *Mech. Time-Dependent Mater.* **2010**, *14*, 219−241.
(40) Yakacki, C. M.; Shandas, R.; Safranski, D.; Ortega, A. M.; Sassaman, K.; Gall, K. Strong, Tailored, Biocompatible Shape-Memory Polymer Networks. *Adv. Funct. Mater.* **2008**, *18*, 2428−2435.
(41) Arrieta, J. S.; Diani, J.; Gilormini, P. Cyclic and Monotonic Testing of Free and Constrained Recovery Properties of a Chemically Crosslinked Acrylate. *J. Appl. Polym. Sci.* **2014**, *131*, 39813.
(42) Santiago, D.; Fabregat-Sanjuan, A.; Ferrando, F.; De la Flor, S. Improving of Mechanical and Shape-Memory Properties in Hyperbranched Epoxy Shape-Memory Polymers. *Shape Mem. Superelasticity* **2016**, *2*, 239−246.
(43) Dutta, R.; Renshaw, D.; Chen, C.; Liang, D. Machine Learning Based Approach for Shape Memory Polymer Behavioural Characterization. *Array* **2020**, *7*, 100036.
(44) Lin, T.-S.; Coley, C. W.; Mochigase, H.; Beech, H. K.; Wang, W.; Wang, Z.; Woods, E.; Craig, S. L.; Johnson, J. A.; Kalow, J. A.; Jensen, K. F.; Olsen, B. D. BigSMILES: A Structurally-Based Line Notation for Describing Macromolecules. *ACS Cent. Sci.* **2019**, *5*, 1523−1531.
(45) Barszczewska-Rybarek, I. M.; Korytkowska-Wałach, A.; Kurcok, M.; Chladek, G.; Kasperski, J. DMA Analysis of the Structure of Crosslinked Poly(Methyl Methacrylate)S. *Acta Bioeng. Biomech.* **2017**, *19*, 47−53.
(46) Lee, J.-W.; Park, W. B.; Do Lee, B.; Kim, S.; Goo, N. H.; Sohn, K.-S. Dirty Engineering Data-Driven Inverse Prediction Machine Learning Model. *Sci. Rep.* **2020**, *10*, 20443.
(47) Gómez-Bombarelli, R.; Wei, J. N.; Duvenaud, D.; Hernández-Lobato, J. M.; Sánchez-Lengeling, B.; Sheberla, D.; Aguilera-Iparraguirre, J.; Hirzel, T. D.; Adams, R. P.; Aspuru-Guzik, A. Automatic Chemical Design Using a Data-Driven Continuous Representation of Molecules. *ACS Cent. Sci.* **2018**, *4*, 268−276.
(48) Miccio, L. A.; Schwartz, G. A. From Chemical Structure to Quantitative Polymer Properties Prediction through Convolutional Neural Networks. *Polymer* **2020**, *193*, 122341.
(49) Wu, S.; Kondo, Y.; Kakimoto, M.-a.; Yang, B.; Yamada, H.; Kuwajima, I.; Lambard, G.; Hongo, K.; Xu, Y.; Shiomi, J.; Schick, C.; Morikawa, J.; Yoshida, R. Machine-Learning-Assisted Discovery of Polymers with High Thermal Conductivity Using a Molecular Design Algorithm. *npj Comput. Mater.* **2019**, *5*, 66.
(50) Lakhera, N.; Yakacki, C. M.; Nguyen, T. D.; Frick, C. P. Partially Constrained Recovery of (Meth)Acrylate Shape-Memory Polymer Networks. *J. Appl. Polym. Sci.* **2012**, *126*, 72−82.
(51) Lu, H.; Huang, W. M. On the Origin of the Vogel-Fulcher-Tammann Law in the Thermo-Responsive Shape Memory Effect of Amorphous Polymers. *Smart Mater. Struct.* **2013**, *22*, 105021.
(52) Kingma, D. P.; Welling, M. Auto-Encoding Variational Bayes. *Proceedings of the 2nd International Conference on Learning Representations (ICLR)*, 2013; pp 1−14.
(53) Samanta, S.; O'Hagan, S.; Swainston, N.; Roberts, T. J.; Kell, D. B. VAE-Sim: A Novel Molecular Similarity Measure Based on a Variational Autoencoder. *Molecules* **2020**, *25*, 1−16.
(54) Samanta, B.; De, A.; Jana, G.; Chattaraj, P. K.; Ganguly, N.; Rodriguez, M. G. NEVAE: A Deep Generative Model for Molecular Graphs. *J. Mach. Learn. Res.* **2020**, *33*, 1110−1117.
(55) Griffiths, R.-R.; Hernández-Lobato, J. M. Constrained Bayesian Optimization for Automatic Chemical Design Using Variational Autoencoders. *Chem. Sci.* **2020**, *11*, 577−586.
(56) Hirohara, M.; Saito, Y.; Koda, Y.; Sato, K.; Sakakibara, Y. Convolutional Neural Network Based on SMILES Representation of Compounds for Detecting Chemical Motif. *BMC Bioinf.* **2018**, *19*, 526.
(57) Matyushin, D. D.; Sholokhova, A. Y.; Buryak, A. K. A Deep Convolutional Neural Network for the Estimation of Gas Chromatographic Retention Indices. *J. Chromatogr. A* **2019**, *1607*, 460395.
(58) Zheng, H.; Fu, J.; Mei, T.; Luo, J. Learning Multi-Attention Convolutional Neural Network for Fine-Grained Image Recognition. *Proceedings of the IEEE International Conference on Computer Vision*, 2017; pp 5219−5227.
(59) Wei, X.; Xie, C.; Wu, J. Mask-CNN: Localizing Parts and Selecting Descriptors for Fine-Grained Image Recognition. **2016**. arXiv:1605.06878.
(60) Yan, C.; Li, X.; Li, G. A New Action Recognition Framework for Video Highlights Summarization in Sporting Events. *International Conference on Computer Science & Education*; IEEE: Lancaster, 2021.
(61) Tsunoda, T.; Komori, Y.; Matsugu, M.; Harada, T. Football Action Recognition Using Hierarchical LSTM. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2017; pp 99−107.
(62) Gammulle, H.; Denman, S.; Sridharan, S.; Fookes, C. Two Stream LSTM: A Deep Fusion Framework for Human Action Recognition. *Proceedings—2017 IEEE Winter Conference on Applications of Computer Vision, WACV 2017*, 2017; pp 177−186.
(63) Wu, E.; Koike, H. FuturePong: Real-Time Table Tennis Trajectory Forecasting Using Pose Prediction Network. *CHI Conference on Human Factors in Computing Systems—Proceedings*, 2020.
(64) Gaulton, A.; Hersey, A.; Nowotka, M.; Bento, A. P.; Chambers, J.; Mendez, D.; Mutowo, P.; Atkinson, F.; Bellis, L. J.; Cibrián-Uhalte, E.; Davies, M.; Dedman, N.; Karlsson, A.; Magariños, M. P.; Overington, J. P.; Papadatos, G.; Smit, I.; Leach, A. R. The ChEMBL Database in 2017. *Nucleic Acids Res.* **2017**, *45*, D945−D954.
(65) Asperti, A. Sparsity in Variational Autoencoders. **2018**. arXiv:1812.07238.
(66) Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J. Chem. Inf. Model.* **2010**, *50*, 742−754.
(67) Lu, H.; Liu, Y.; Leng, J.; Du, S. Qualitative Separation of the Physical Swelling Effect on the Recovery Behavior of Shape Memory Polymer. *Eur. Polym. J.* **2010**, *46*, 1908−1914.
(68) Lu, H.; Du, S. A Phenomenological Thermodynamic Model for the Chemo-Responsive Shape Memory Effect in Polymers Based on Flory-Huggins Solution Theory. *Polym. Chem.* **2014**, *5*, 1155−1162.
(69) Magzoub, M.; Salehi, S.; Li, G.; Fan, J.; Teodoriu, C. Loss Circulation Prevention in Geothermal Drilling by Shape Memory Polymer. *Geothermics* **2021**, *89*, 101943.

https://doi.org/10.1021/acsami.1c20947
ACS Appl. Mater. Interfaces 2021, 13, 60508−60521