


Article

# SCORN: Sinter Composition Optimization with Regressive Convolutional Neural Network

Junhui Li<sup>1</sup>, Liangdong Guo<sup>1</sup> and Youshan Zhang<sup>2,\*</sup> 

<sup>1</sup> Department of Mathematics, School of Science, University of Science and Technology, Anshan 114051, China; junhui\_lee@foxmail.com (J.L.); ldguo@ustl.edu.cn (L.G.)

<sup>2</sup> Department of Clinical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY 14853, USA

\* Correspondence: yz945@cornell.edu

**Abstract:** Sinter composition optimization is an important process of iron and steel companies. To increase companies' profits, they often rely on innovative technology or the workers' operating experience to improve final productions. However, the former is costly because of patents, and the latter is error-prone. In addition, traditional linear programming optimization methods of sinter compositions are inefficient in the face of large-scale problems and complex nonlinear problems. In this paper, we are the first to propose a regressive convolutional neural network (RCNN) approach for the sinter composition optimization (SCORN). Our SCORN is a single input and multiple outputs regression model. Sinter plant production is used as the input of the SCORN model, and the outputs are the optimized sintering compositions. The SCORN model can predict the optimal sintering compositions to reduce the input of raw materials consumption to save costs and increase profits. By constructing a new neural network structure, the RCNN model is trained to increase its feature extraction capability for sintering production. The SCORN model has a better performance compared with several regressive approaches. The practical application of this predictive model can not only formulate corresponding production plans without feeding materials but also give better input parameters of sintered raw materials during the sintering process.



**Citation:** Li, J.; Guo, L.; Zhang, Y. SCORN: Sinter Composition Optimization with Regressive Convolutional Neural Network. *Solids* **2022**, *3*, 416–429. <https://doi.org/10.3390/solids3030029>

Academic Editor: John D. Clayton

Received: 1 June 2022

Accepted: 9 July 2022

Published: 12 July 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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 (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** sinter; composition optimization; regressive CNN; multiple output regression

## 1. Introduction

Sinter has always been an important part of the steel-making process in a sintering plant. Sintering technology is a complex thermo-chemical and energy-intensive process, and the price of its raw material—iron ore—has always been high. As a result, how to control costs and improve profits are the core issues that can affect the survival of sinter plant enterprises [1,2]. Research on sinter composition optimization is an extremely important field of sinter mineralogy, and the quality of ingredients affects the final sinter quality. In most practical cases, the proportion of sinter ore is limited by manual experience, which is subjective. It is also difficult to obtain the optimal material ratio due to the contradiction among the constraints [3]. The sintering process modeling method and linear programming are widely used to address this challenge. However, one problem is that there are many nonlinear factors that need to be considered in the batch optimization model [4]. With the development of the research on the sintering process, mineral varieties are increasing. The number of chemical composition control projects is also increasing. Because sintering mixtures are complex, it is difficult to change one parameter independently of others, and the introduction of new parameters into the optimization model can simultaneously change all parameters [5], which is time-consuming and tedious to calculate.

In this paper, we propose a new sinter compositions optimization model using a regressive convolutional neural network (SCORN). We develop a new regressive convolutional neural network (RCNN) structure from given datasets to obtain the optimal sintering material compositions. By using the production history data of a mining company in China,

our SCORN model can predict the optimal sinter compositions. The proposed SCORN model includes feature extraction and prediction modules. Unlike conventional machine learning tools and CNNs, the input is only a single number (production), and the output is the corresponding chemical indexes of the final sinter. This model is a multi-output model, and the final sintering ratio scheme consists of multiple indicators. Our contributions are two-fold:

1. We are the first to develop a regressive convolutional neural network for the sinter composition optimization problem. In our SCORN model, the input is the single final sintering production, and outputs are the corresponding chemical compositions of the sintered product. SCORN is a single input and multiple outputs RCNN model.
2. We have collected sinter production and its burdening compositions from sintering machines in one sintering plant in China. Experimental results indicate that our SCORN model can produce an optimal sinter burdening ratio given a target production. SCORN also achieves higher performance than several regressive approaches.

Our paper aims to extract features from sinter production data to predict optimal sinter compositions of that production. Because of the single input data, the RCNN architecture needs to be efficient and accurate to extract the key features. Therefore, linear programming and intelligent optimization algorithms for solving multivariate input problems are ineffective in our problem.

The rest of the paper is organized as follows. Section 2 provides an overview of related work for sinter optimization. The description of the sintering process and characteristic indexes are mentioned and summarized in Section 3. Section 4 provides details about the proposed methods, including structure and evaluation methods. Section 5 provides a detailed evaluation of the SCORN with a solid comparison with other regressive methods on the same sinter datasets. This section is further divided into subsections to describe the details of the dataset and the experimental setup for the traditional approaches. Section 6 discusses the model, including advantages and disadvantages, as well as practical applications and extensions. Finally, Section 7 concludes the paper and outlines directions for possible future work.

## 2. Related Work

In the past few decades, scholars have carried out many research methods to optimize various iron ore sinter indicators to improve the sintering performance and reduce the cost.

### 2.1. Mathematical Statistical Models

Many studies have attempted to address the question of predicting sinter quality, properties, and productivity. Many sinter models have been constructed based on mathematical-statistical methods. Eugene et al. [6] presented a mathematical modeling method to predict sinter properties. This method reflected the variation in sinter properties using explanatory variables and optimized different iron ore blends to produce target sinter characteristics. Zhang et al. [7] developed an unsteady two-dimensional mathematical model for the iron ore sintering process and predicted sinter yield and strength by the method of numerical simulation. In view of the large time lag in the detection of sinter, Li et al. [8] verified the relationship between the chemical compositions of the sintering raw material and the physical and metallurgical properties of the sinter through correlation analysis. However, the aforementioned mathematical models are mainly optimized from the aspects of sintering process parameters and properties and do not consider many other factors in the sintering process. Due to the difference between ideal models and actual processes, they are difficult to apply to industrial processes.

### 2.2. Machine Learning

Various machine learning tools and intelligent optimization algorithms are increasingly used in the sinter process research. Support vector machines, BP neural network models, and general regression neural network [9] models have been applied as prediction models

for basic sintering characteristics and sinter quality of mixed iron ore. Arghya et al. [10] associated the sinter plant process parameters with required mechanical properties and microstructure to obtain higher productivity with the help of ANN and genetic algorithms. Kunnunen et al. [11] shed light on how neural networks were used to model and optimize physical indexes of sinter. Yuan et al. [12] applied a deep belief network algorithm to predict the secondary chemical composition of the sinter by analyzing the technology mechanism and characteristics of the sintering process. Machine learning methods have been widely used in the field of the sintering process to optimize the relevant indicators. Most deep neural networks address the detection [13–15] and classification [16] problems in the sintering process. Frei et al. [17] proposed a novel deep learning-based method for the pixel-perfect detection and the measurement of partially sintered particles. It is difficult for shallow learning algorithms to effectively represent complex nonlinear functions when the number of given samples is limited. The generalization ability is also limited, which affects the prediction results of the sinter composition optimization problem.

Deep learning is a branch of machine learning and relies on a large amount of data to build models that estimate the patterns of the data. Over the past two decades, CNNs have relied on the hidden layer structure to automatically extract deep features, which achieved promising results in a wide range of vision applications and domains such as image denoising, image detection, and classification [18]. Le and Ho [19] presented a novel method to predict DNA 6 mA sites from the cross-species genome based on deep transformers architecture and CNN with DNA sequence as input. Le and Nguyen [20] proposed a method to identify FMN binding sites in electron transport chains using a 2-D CNN constructed from position-specific scoring matrices (PSSM). The proposed method can also facilitate the application of deep learning to deal with various problems in bioinformatics and computational biology. Aziz et al. [21] developed a new technique of Channel Boosted Convolutional Neural Network (CB-CNN) to classify breast cancer mitotic nuclei. This method improves the generalization of a CNN by making the feature space more versatile and flexible.

### 2.3. Sinter Compositions

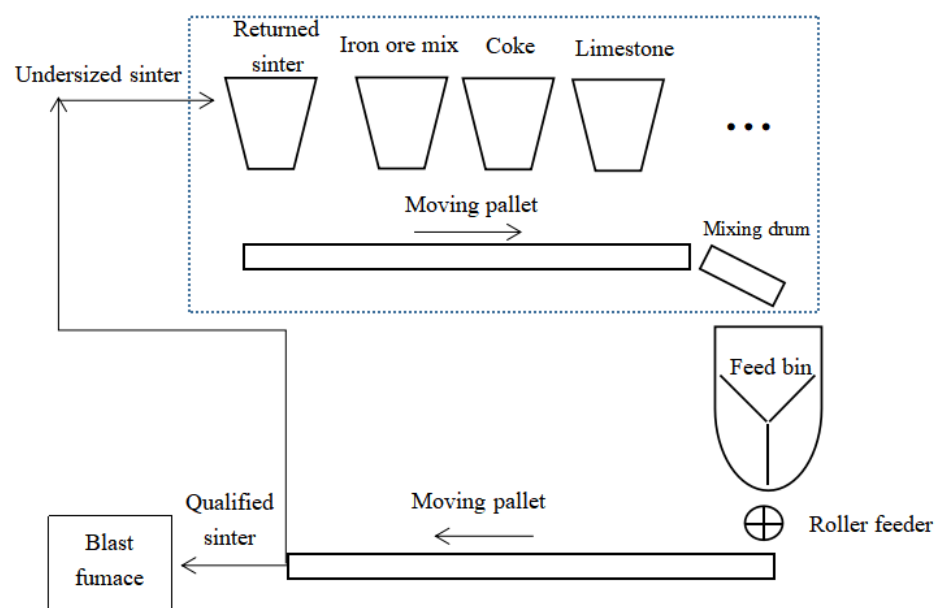
Existing works try to optimize the sinter composition and reduce the cost of the sintering process. Efforts are being made to resolve the proportioning issues associated with the sintering process. Based on the micro-sintering experiment [22], the principle of ore blending is put forward according to its high-temperature characteristics. Then the ore blending is optimized. Linear programming (LP) and nonlinear programming (NLP) methods are also commonly used for evolutionary optimization of blast furnace charging ratios and operating parameters [23–26]. Most of these methods used the cost as an objective function, but in practice, the optimization objective is often multi-fold, making it challenging to meet the requirements of the sintering process. Liu et al. [27] proposed a real-time monitoring model and advanced prediction of sinter composition based on a DNN and LSTM regression. Taking the lowest cost of sinter as the objective function, Wang and Hu [4] established a comprehensive optimization model of sinter batching and solved it with the particle swarm algorithm (PSO). Dai and Zhen [3] established a genetic chickens hybrid algorithm based on linear programming, which is used in the first and second compositions optimization of the sintering process. Wu et al. [28] developed an intelligent integrated optimization system (IIOS) for the sintering ratio step to find the best feasible proportion regimen. The optimal burdening ratio method using intelligent optimization algorithms has been extensively studied, including SA (simulated annealing algorithms), EA (evolutionary algorithms), PSO algorithms, ACA (ant colony algorithms), etc. However, they all have a common problem. These algorithms converged quickly at first but then became slower, making it easy to obtain the locally optimal solution [29–31].

### 3. Sintering Process and Characteristic Indexes

This section briefly describes the sintering process and explains the physical and chemical characteristic indexes for sintered final products.

#### 3.1. Description of Sintering Process

The entire sintering preparation process is complex, mainly including three steps: batching, mixing, and sintering. In the sintering batching stage, the chemical raw materials of sintered ore and other materials are mixed in a certain proportion. After the mixing stage, contents are evenly mixed with water and then sent to the sinter machine to generate sintered ore. The sintering process undergoes complex physical and chemical changes, and the entire process can take up to two hours or more [32]. Figure 1 shows the main material flows in the sintering process.



**Figure 1.** Material flows in the sintering process

#### 3.2. Sinter Characteristic Indexes

The indexes of iron ore characteristics are selected from the following two aspects.

##### 3.2.1. Chemical Index

The chemical index mainly consists of two parts. Firstly, the chemical composition part of sinter generally includes TFe, SiO<sub>2</sub>, MgO, Al<sub>2</sub>O<sub>3</sub>, CaO, S, and FeO. Secondly, other indexes are Ro and total iron ore. Ro is expressed as the ratio of calcium oxide content to silica content in the sinter. The total amount of sinter represents the sum of the total chemical components of the sinter.

##### 3.2.2. Physical Index

Screening is defined as the percentage of sintered ore smaller than the standard specified particle size (−5 mm) in the total weight of the sample after the sample is screened. The drum index is defined as the percentage of the weight of a sample with a particle size larger than the specified standard to the total weight of the sample. Table 1 shows the characteristic information iron ore, which is mainly composed of chemical indicators.

**Table 1.** Statistics for sintering compositions

| Factors                      | Field                          | Unit | Average/Year | Change Range |
|------------------------------|--------------------------------|------|--------------|--------------|
| Chemical compositions        | TFe                            | %    | 56.2         | 53.9–58.6    |
|                              | FeO                            | %    | 8.8          | 5.5–12.3     |
|                              | SiO <sub>2</sub>               | %    | 5.7          | 4.7–6.4      |
|                              | CaO                            | %    | 11.6         | 9.7–13.7     |
|                              | RO                             | -    | 2            | 1.6–2.3      |
|                              | MgO                            | %    | 1.5          | 0.9–2.1      |
|                              | S                              | %    | 0.026        | 0.002–0.063  |
|                              | Al <sub>2</sub> O <sub>3</sub> | %    | 1.23         | 0.19–1.98    |
| Total iron ore in the sinter |                                | -    | 98.35        | 96.88–99.80  |

## 4. Methods

### 4.1. Motivation

In the pre-iron process for iron and steel enterprises, an efficient and accurate grasping of the current sinter composition is of great significance for guiding blast furnace production. The metallogenetic process of the sintering mixture is complex. It is difficult to accurately obtain the optimal sintering compositions corresponding to the mixture through mechanism calculation. Statistics-based machine learning methods can rely on large-scale data to obtain a reliable prediction model. The feature extraction depends more on the hidden layer model and is better at processing high-dimensional data. The quality of the sinter is closely related to batching, process state, and operating parameters. Traditionally, the appropriate ratio for sintering production is determined by chemical principles and a large number of experiments. Then linear programming or intelligent optimization algorithms are used to optimize it. Under industrial conditions, where production depends on the used raw materials, there is no simple answer to the question of how a certain value is optimized. The purpose of this study was to produce refined knowledge that would assist in the control of the sinter composition value when the production is determined. In comparison to conventional ANNs, RCNNs apply a largely increased number of layers, which can extract complicated features [33]. Our research problem is an optimization task, and we aim to optimize the sinter compositions using a RCNN given the sinter production data.

### 4.2. Problem

In the sinter plant, production increase often only relies on technological innovation or skilled operation. However, it is not always reliable to depend on the experience of operators. The results obtained by each person using these methods are not consistent, and it is not easy to accurately control the burdening ratio of the sinter. The raw materials in sintering production consist of many different compositions, and each composition may have a mutual influence or correlation. Therefore, a sinter burdening ratio optimization model based on RCNN is proposed to solve the problems. In our case, there is an unknown relationship between target production and the chemical composition of the sinter: the input of this model is the target production, and the output is the chemical index of the sinter.

### 4.3. Notations

In the sinter composition optimization problem, given the  $N$  input target sintering productions  $X = \{x_n\}_{n=1}^N \in \mathbb{R}^{N \times 1}$ , its outputs are the optimized sinter compositions  $Y = \{y_n\}_{n=1}^N \in \mathbb{R}^{N \times D}$ , where  $D$  is different indexes that are mentioned in Section 3.2.1. Each instance is characterized by an input sintering production  $x_n$  and an output sintering composition  $y_n \in \mathbb{R}^{1 \times D}$ . The objective function of our proposed SCORN model is to train a regressive convolution neural network ( $\mathcal{R}$ ) to accurately predict optimal sintering composition given any input production as follows:

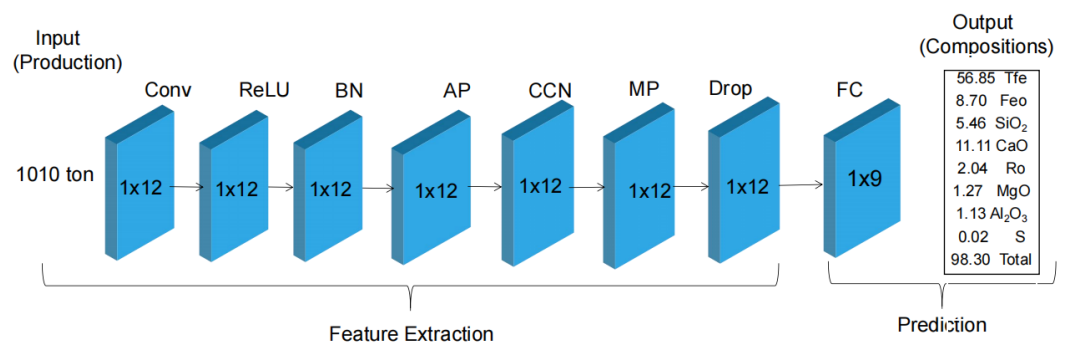
$$X = \{x_n\}_{n=1}^N \xrightarrow{\mathcal{R}} Y = \{y_n\}_{n=1}^N. \quad (1)$$

#### 4.3.1. Architecture

Our proposed SCORN model consists of two major modules [34].

1. **Feature Extraction.** The feature extraction module extracts features from the simple numerical target production for the second module. One advantage of the RCNN architecture is that the layers are easily interchangeable, which greatly facilitates transfer learning between layers [17].
2. **Prediction.** This block takes the extracted features from the previous module and feeds them to a fully connected (FC) layer for regression prediction.

To overcome the smaller number of features of the input layer (production has the size of  $1 \times 1$ ) in the first module, we need to design an appropriate network architecture for extracting better feature representations to model the relationship between the production and nonlinear indexes. Notably, the input of the model is sintering production, and the output sintering compositions come from the final connected regression layer. To achieve better accuracy, the feature extraction structure may be used multiple times. Figure 2 shows an example of a sinter composition prediction of the SCORN model. The final few layers can reflect completed sinter compositions. With more features extracted in the feature extraction module, we can easily build the relationship between the model and the predicted sinter compositions.



**Figure 2.** The architecture of our proposed SCORN for predicting the compositions of sinter at “production” 1010 ton (Conv: convolution, ReLU: rectified linear units, BN: Batch normalization, AP: average pooling, CCN: cross-channel normalization, MP: max pooling, Drop: dropout and FC: fully connected. The number of features of each layer is represented in the middle of the graph of each layer).

Normally, CNN consists of a sequence of layers, including convolutional layers, pooling layers, and fully connected layers. Each convolutional layer typically has two stages. In the first stage, the layer performs the convolution operation, which results in linear activation. In the next stage, a nonlinear activation function is applied to each linear activation. Each feature extraction module [35] has seven layers (convolution (Conv), rectified linear units (ReLU), batch normalization (BN), average pooling (AP), cross-channel normalization (CCN), dropout (Drop) and max pooling (MP)). SCORN model can extract features from the simple numerical target production and feed them into a fully connected (FC) layer for regressive sinter composition prediction.

In the SCORN model, we employ the Conv layer to generate more features from the previous layer (e.g., the first Conv layer has the filter size of [1, 1], number of filters: 12, stride size of [1, 1] and zero padding. Hence, the final output size is  $1 \times 12$ ). The ReLU layer reduces the number of epochs to achieve better training error rates than traditional tanh units. The normalization layer increases the generalization ability and reduces the error rate. In addition, ReLU and normalization layers do not change the size of the feature map. The pooling layer aggregates the outputs of adjacent pooling units. The dropout (Drop)



layer randomly sets input elements to zero to prevent overfitting. The loss function of the last regression layer is the same as our error function. One of the most obvious advantages of the model is that more features can be extracted from the feature extraction model. By extracting more features in the model, we can easily establish the relationship between the model and the predicted components [36]. Therefore, the SCORN model can predict composition at an arbitrary production.

#### 4.3.2. Loss Function

The half sum-of-squared errors in Equation (2) has been employed as an indicator of the discrepancy between the actual  $y_n$  and the predicted output  $y'_n$ . By reducing the error between the actual and the predicted value, the SCORN model can predict the sintering compositions.

$$E = \frac{1}{2} \sum_{n=1}^N [y_n - y'_n]^2 \quad (2)$$

#### 4.4. Model Evaluation

To illustrate the significance of the SCORN model, we focus not only on the fitting effect of the model but also on the error values between the predicted value and the real value. Therefore, the extended  $R^2$ , root mean square error (RMSE), and mean absolute error (MAE) were used to evaluate the model, as in Equations (3)–(5).

$$R^2 = 1 - \frac{\text{Unexplained variation}}{\text{Total variation}} = 1 - \frac{\sum_{n=1}^N S_{\text{residual}}}{\sum_{n=1}^N S_{\text{total}}}, \text{ where} \quad (3)$$

$$S_{\text{residual}} = \sum_{d=1}^D (y_{nd} - y'_{nd})^2 \quad S_{\text{total}} = \sum_{d=1}^D (y_{nd} - \bar{y}_d)^2$$

$$\text{RMSE} = \frac{1}{D} \sum_{d=1}^D \sqrt{\frac{1}{N} \sum_{n=1}^N [y_{nd} - y'_{nd}]^2} \quad (4)$$

$$\text{MAE} = \frac{1}{ND} \sum_{n=1}^N \sum_{d=1}^D |y_{nd} - y'_{nd}| \quad (5)$$

In these formulas,  $y_{nd}$  is the actual value of the  $d$ th data point, and  $y'_{nd}$  is the predicted value.  $N$  is the number of samples in the sinter composition.  $D$  is the number of composition indexes of the sintering process. The  $R^2$  statistic has been shown to be a useful indicator of the significance of the model's performance [37]. Therefore, our unknown relationship regression is fitted with an extended  $R^2$  statistic. The range of the  $R^2$  statistic is between  $[0, 1]$ ; the higher the value of  $R^2$ , the more variation the model explains, and the better the model fits the sinter composition. In addition, the smaller RMSE and MAE, the better the model is. We used a five-fold cross-validation method to evaluate model performance.

Hypothesis tests: We test a hypothesis to show the significance of predicted and true sinter compositions. The null hypothesis is  $H_0$ : there is no significant difference between predicted sinter compositions and original sinter compositions (they come from the same distribution). We perform two-sample t-tests to calculate the  $p$ -values. Since each prediction will have a  $p$ -value, we compute the mean  $p$ -value of the whole dataset.

## 5. Experimental Setups

We evaluated our model on a sintering dataset and provided a detailed comparison with six regression methods. This section also provides a detailed description of the dataset and its evaluation settings.

## 5.1. Datasets Description

### 5.1.1. Sinter Datasets

One of the most important aspects of any machine learning method is having input and output data from reliable sources. Usually, there is a seasonal variation in the input parameters, such as the percentage fluctuation of MaO in iron ore, which is usually lower during the rainy season. To predict the sinter composition ratio, the chemical indexes from the database of a sintering plant in China were collected. The time of data spans from January 2017 to December 2018. The data of the whole production line can be classified into mainly two categories: chemical index and physical index, as mentioned in Section 3.2. For the sinter data, 12 manual samplings and analyses are performed daily. Daily data from a period of two years were used in our data-driven modeling. The period yielded a set of 7803 valid observations of the model. The statistics of sintering compositions can be seen in Table 1. Finally, our model was also developed to correlate nine sinter compositions as the output variable and the sinter production, described as input variables.

### 5.1.2. More Validation Datasets

We also used three external datasets (Pentagon, Corpus Callosum, and Mandible shape; the details of these datasets can be found in [35]) to validate our model. We further compared our model with a geodesic regression and ShapeNet models [35].

## 5.2. Implementation Details

We compared the results of regressive methods with the mentioned datasets via MATLAB software and Python using an Intel(R) Core(TM) i5-10500 CPU. We used 6242 (70% of dataset) as the training set and the rest 1561 (30% data) as the test set. We compared the predicted results of our SCORN with the other six baseline methods (DecisionTree [38], RForest [39], KNN [40], LS [41], MLP [42], SVR [43]). Our SCORN model's structure finally has ten layers. The number of composition indexes of the sintering process  $D$  is nine. We chose sgdm as our optimization function. The maximum number of iterations was set as 300, and the initial learning rate was set as 0.0005. The running time of training our model is 350 s, and the inference time is less than 0.1 s. To train the DecisionTree model, we used the default parameters from Python's Scikit-Learn module. For SVR, the algorithm does not support multiple outputs for regression problems, and we implemented multi-objective support vector regression via a correlation regression chain [44,45]. We used the RBF (radial basis function kernel) kernel, and other parameters were set as default values. For the MLP model, the training was started with a simple 20-50-100 structure hidden layer, and we chose tanh as our activation function. The maximum number of iterations was set as 100, and the penalty function was set as 0.0001. Different regressive methods were developed. Each method was started based on the same datasets.

## 5.3. Results

In this section, we provide a detailed comparison with six conventional methods. The significant analyses demonstrated the applicability and goodness of our model.

### 5.3.1. The Traditional Methods Used for Comparison

This part summarizes other used regression models that are compared with our SCORN model. Many machine learning algorithms are designed to predict a single numerical value, referred to as a single output regression model. However, we can also encounter many multi-output regression problems in real life. Multi-output regression aims to learn a mapping between a single or multivariate input space and a multivariate output space [41].

1. **Least Square.** Least squares is a mathematical optimization technique that finds the best functional match for the data by minimizing the sum of squared errors.
2. **KNN.** The nearest-neighbor technique is a well-known and studied technique in statistical learning theory [40]. In essence, the method consists of constructing estimators



by averaging the properties of training events of similar characteristics to those of a test event to be classified or whose properties need to be inferred.

3. **RandomForest.** A random forest algorithm is an ensemble approach that relies on CART models [39].
4. **Decision Tree.** In a decision tree model, an empirical tree represents a segmentation of data, which is created by applying a series of simple rules. These models generate a set of rules that can be used for prediction through the repetitive process of splitting [38].
5. **Multilayer Perceptron.** MLPs learn a mapping function from the input space to the target space [42]. Generally, there are three basic layers in the structure of MLPs, the input layer, the number of hidden layers, and the output layers. The three-layer MLP consists of one input node, three hidden layers with [20, 50, 100] hidden nodes, and nine output nodes in each joint.
6. **SVR.** Support vector regression (SVR) works on the principle of structural risk minimization (SRM) from statistical learning theory. The core idea of the SRM theory is to arrive at a hypothesis  $h$ , which can yield the lowest true error for the unseen and random sample testing data [43].

### 5.3.2. Composition Predictions

After training the SCORN model with the sinter plant training set, we applied the model to predict the sinter composition of the test set. The training curve and validation curve of the trained network structure is shown in Figure 3. The comparison results of actual and predicted compositions are shown in Table 2. We enumerated the predicted values of six groups of samples and their corresponding true values. Compared with the actual composition with the SCORN predicted composition, the predicted value is close to the actual component, which indicates that the SCORN model has a good prediction effect. Figure 4 shows the detailed comparison between the prediction and the original value of sinter composition  $T_{Fe}$  based on the SCORN model. Most of the predicted values are close to the original values. Both the predicted value and the original value fluctuate within the same numerical range, which shows that the SCORN model has a high generalization ability.

**Table 2.** Comparison of the actual composition and the predicted composition using our SCORN model (# means the actual number, and P# represents the predicted number).

| Compositions                     | # 1   | P# 1  | # 2   | P# 2  | # 3   | P# 3  | # 4   | P# 4  | # 5   | P# 5  | # 6   | P# 6  |
|----------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| TFe %                            | 56.98 | 56.86 | 57.22 | 56.94 | 57.08 | 56.94 | 56.90 | 56.95 | 56.91 | 56.85 | 56.82 | 56.85 |
| Feo %                            | 7.7   | 8.7   | 8.9   | 8.65  | 9     | 8.65  | 8.4   | 8.65  | 8.40  | 8.70  | 8.90  | 8.69  |
| SiO <sub>2</sub> %               | 5.52  | 5.46  | 5.39  | 5.46  | 5.47  | 5.47  | 5.38  | 5.47  | 5.39  | 5.46  | 5.54  | 5.46  |
| CaO %                            | 11.03 | 11.11 | 10.85 | 11.01 | 11.05 | 11.01 | 10.85 | 11.01 | 11.22 | 11.11 | 11.52 | 11.10 |
| Ro                               | 2     | 2.04  | 2.01  | 2.02  | 2.02  | 2.02  | 2.02  | 2.02  | 2.08  | 2.04  | 2.08  | 2.03  |
| MgO %                            | 1.55  | 1.27  | 1.52  | 1.20  | 1.49  | 1.20  | 1.5   | 1.20  | 1.61  | 1.27  | 1.53  | 1.26  |
| S %                              | 0.86  | 1.12  | 0.85  | 1.13  | 0.88  | 1.13  | 0.89  | 1.13  | 0.86  | 1.12  | 0.86  | 1.12  |
| Al <sub>2</sub> O <sub>3</sub> % | 0.032 | 0.022 | 0.024 | 0.021 | 0.026 | 0.021 | 0.027 | 0.021 | 0.028 | 0.022 | 0.028 | 0.023 |
| Total iron ore                   | 98.87 | 98.31 | 98.73 | 98.27 | 98.77 | 98.27 | 98.30 | 98.27 | 98.81 | 98.30 | 98.99 | 98.29 |

### 5.3.3. Significance Analysis

After predicting the sinter composition of the test set, we calculated the statistical significance of SCORN and each comparison method that is described in Section 5.3.1. Table 3 shows the values of  $R^2$ , RMSE and MAE, as given in Equations (3)–(5), of the training set using the five-fold cross-validation method. The  $R^2$  score of different methods is close to 1, which shows that the fitting degree of the model is good. We also reported the uncertainty of all models. Except for the SVR model, the RMSE and MAE of the SCORN and other compared models are generally low. In addition, the  $R^2$  score and RMSE of the SCORN model are better than those of other models. MAE of the SCORN model is also close to the best value. The higher  $R$ -value of the SCORN model shows that the prediction

performance of the SCORN model is better than the other traditional models. This result shows that our model can be practically applied in situations where a large amount of data is available. Similarly, RMSE, the standard deviation of residuals, is smaller than that of other regressive models. It shows that residuals are dispersed in a narrower range in the case of the SCORN model compared to the other regressive models. In the present CNN training, a fixed value of the learning rate, 0.0005, was selected. The  $R$ -value may be further increased if the dynamic learning rate is used [46].

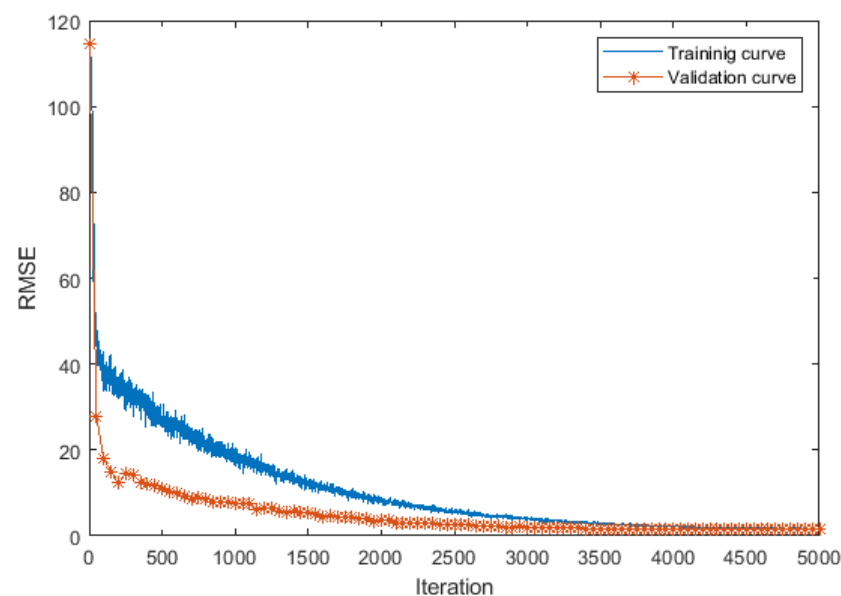
The mean  $p$ -value of the whole dataset is 0.995. All results are from two-sample  $t$ -tests and cannot reject the null hypothesis, which implies that the predicting sintering compositions are similar to the true sintering compositions in which the predicted values almost recover the original values. Table 4 compares the  $R^2$  statistic of SCORN with geodesic regression and ShapeNet models. The  $R^2$  values of three datasets from our SCORN are much larger than those of other models. The lower values indicate that shape variability is not well modeled by the geodesic regression model. Therefore, our SCORN model shows higher effectiveness in predicting the shapes of three different validation datasets given a single input.

**Table 3.** Comparison of the SCORN models and the different methods (the variances of  $R^2$  values are negligible because of their small values).

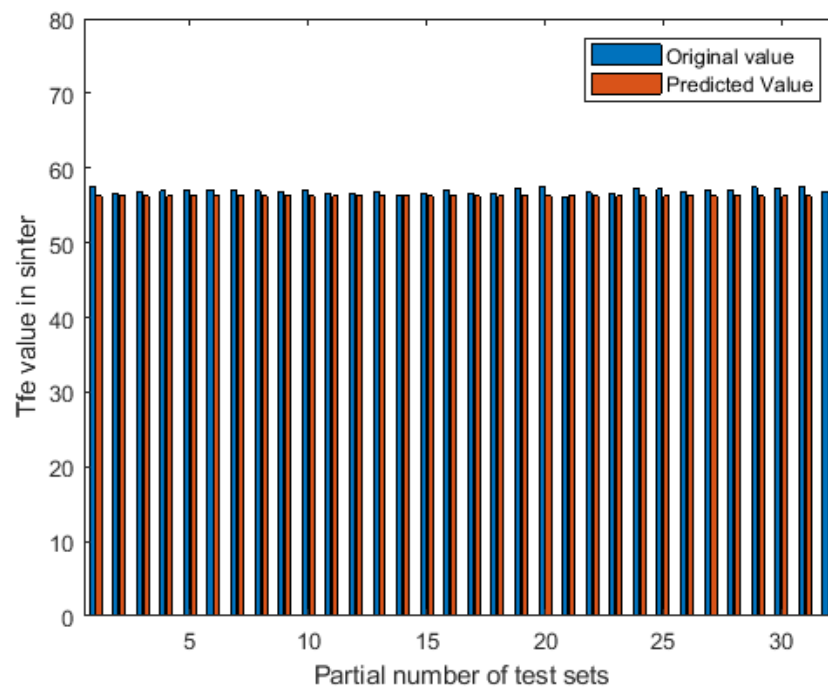
| Method | SVR             | KNN             | R Forest        | Decision Tree   | OLS             | MLP             | SCORN           |
|--------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| RMSE   | $3.06 \pm 0.33$ | $0.50 \pm 0.02$ | $0.46 \pm 0.01$ | $0.46 \pm 0.01$ | $0.48 \pm 0.01$ | $0.49 \pm 0.01$ | $0.40 \pm 0.01$ |
| MAE    | $1.18 \pm 0.08$ | $0.33 \pm 0.01$ | $0.31 \pm 0.01$ | $0.31 \pm 0.01$ | $0.33 \pm 0.01$ | $0.34 \pm 0.01$ | $0.33 \pm 0.01$ |
| $R^2$  | 0.9865          | 0.9998          | 0.9998          | 0.9998          | 0.9997          | 0.9997          | 0.9999          |

**Table 4.**  $R^2$  statistic of predicting data.

| Datasets         | Pentagon | Corpus Callosum | Mandible |
|------------------|----------|-----------------|----------|
| Geodesoc regress | 0.0223   | 0.0234          | 0.0873   |
| ShapeNet         | 0.3911   | 0.3854          | 0.1738   |
| SCORN            | 0.9923   | 0.9996          | 0.8191   |



**Figure 3.** Training and validation curve.



**Figure 4.** Comparison of the actual Tfe component and the predicted Tfe component with our SCORN model. The max error is  $\pm 1.4835$ .

#### 5.4. Parameter Analysis

The ablation study of different dropout rates is shown in Table 5. Different dropout rates affect the performance of our SCORN model because different rates can correct errors in other units to help avoid the overfitting problem. Combining Tables 5 and 6, we can find that when the dropout rate is 0.7 and the learning rate is 0.0005, the SCORN model has the best overall performance and high prediction accuracy.

**Table 5.** Ablation study of different dropout rates.

| Evaluation   | drop = 0.5 | drop = 0.6 | drop = 0.7 | drop = 0.8 |
|--------------|------------|------------|------------|------------|
| Traning Loss | 1.5056     | 1.0434     | 1.0056     | 1.0397     |
| Traning RMSE | 1.7353     | 1.4446     | 1.4182     | 1.4420     |

**Table 6.** Ablation study of different learning rates.

| Evaluation   | 0.0001 | 0.0005 | 0.001 |
|--------------|--------|--------|-------|
| Traning Loss | 1.0056 | 0.9723 | 1345  |
| Traning RMSE | 1.4182 | 1.3945 | 51.86 |

## 6. Discussion

In this paper, we exploit the excellent representation learning capability of the deep networks to optimize sinter compositions from the sinter production. We propose a sinter composition optimization model based on an RCNN. From these experiments, we find that the proposed approach can predict the sinter composition changes with a higher  $R^2$  value. One reason is that the network architectures provide enough modeling capacity to encode the sinter chemical composition at each production and generalize it to unseen production. Finally, we note that the model can be easily extended to support more than a single input; natural extensions would include other influential factors such as product class and other indicators. The essential benefit of our proposed model over traditional methods is that our model has better prediction accuracy, which can effectively save the cost of the sintering process.

### 6.1. Applications and Extensions

The technical staff can quickly obtain the optimal raw material ratio using our predicted sintering output. Then the ore mixing structure can be optimized, and the cost can be effectively reduced. In addition, with a more accurate raw material composition, it is beneficial to improve the planning of sintering material scheduling in the sinter plant. Procurement personnel can optimize the plan and cost of iron ore raw materials through a more reasonable economic value assessment of various raw materials. Based on optimization results of sinter composition, it can be further extended to the blast furnace proportioning model. By adding pellets, lump ore, and other related raw materials, a blast furnace batching optimization model can be applied to calculate the optimal raw material ratio of molten iron.

### 6.2. Advantages and Limitations

There are several advantages of our proposed SCORN model. Firstly, our model is only data-driven and can extract key features efficiently and accurately with only a small amount of key input data. Therefore, our model is very convenient for adding new sintering ratio factors. Secondly, the SCORN model can be applied to the sintering process of other types of ores. The sintering composition optimization model can be established without conducting real experiments using raw materials, which saves time and costs. Lastly, the predicted optimized sintering composition of our model meets the premise quality requirements.

One limitation of our work is that not all indicators of sintering granulation characteristics are considered, such as middle size proportion (MSP), average size index (ASI), etc. As for future work, aside from collecting more data, combining our model with pelletizing metrics may improve sinter quality and steel quality. In addition, another disadvantage of our model is that it is not sensitive enough to small changes in sinter production from one time-unit to the next.

## 7. Conclusions

In this paper, we are the first to propose a sinter composition optimization model based on a regressive convolutional neural network. The proposed SCORN model can handle small amounts of data and high-dimensional data. The prediction accuracy of the model is further improved by optimizing the parameters and structure of the RCNN model. Experimental results show that our method performs better than several regressive models. Therefore, our SCORN model is more suitable for predicting the composition of the sintering process in metallurgical enterprises. In the future, we will pay more attention to other physical indexes, the metallurgical properties indexes, and the correlation between the data from the sinter production line. We aim to mine the important parameters that affect the fluctuation of sinter components and build a better model by combining the physical indexes and metallurgical properties indexes.

**Author Contributions:** Conceptualization, Y.Z. and J.L.; methodology, Y.Z.; writing—original draft preparation, J.L.; writing—review and editing, Y.Z. and L.G.; supervision, Y.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Source dataset and code are available at <https://github.com/YoushanZhang/SCORN>.

**Acknowledgments:** The authors thank the sinter company and its employees for their support during the data collection.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Cheng, Z.; Tan, Z.; Guo, Z.; Yang, J.; Wang, Q. Recent progress in sustainable and energy-efficient technologies for sinter production in the iron and steel industry. *Renew. Sustain. Energy Rev.* **2020**, *131*, 110034. [[CrossRef](#)]
2. Naito, M.; Takeda, K.; Matsui, Y. Ironmaking technology for the last 100 years: Deployment to advanced technologies from introduction of technological know-how, and evolution to next-generation process. *ISIJ Int.* **2015**, *55*, 7–35. [[CrossRef](#)]
3. Dai, Z. Sintering Proportioning Optimization: Use LP and GA-CSO. In Proceedings of the 2021 3rd World Symposium on Artificial Intelligence (WSAI), Online, 18–20 June 2021; pp. 34–40.
4. Wang, Y.; Hu, Q. Research and application of optimization method for iron and steel sintering ingredients. In Proceedings of the 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 12–14 October 2018; pp. 1819–1823.
5. Waters, A.G.; Litster, J.D.; Nicol, S.K. A mathematical model for the prediction of granule size distribution for multicomponent sinter feed. *ISIJ Int.* **1989**, *29*, 274–283. [[CrossRef](#)]
6. Donskoi, E.; Manuel, J.R.; Clout, J.M.; Zhang, Y. Mathematical modeling and optimization of iron ore sinter properties. *Isr. J. Chem.* **2007**, *47*, 373–379. [[CrossRef](#)]
7. Zhang, B.; Zhou, J.; Li, M. Prediction of sinter yield and strength in iron ore sintering process by numerical simulation. *Appl. Therm. Eng.* **2018**, *131*, 70–79. [[CrossRef](#)]
8. Li, Y.; Zhang, Q.; Zhu, Y.; Yang, A.; Liu, W.; Zhao, X.; Ren, X.; Feng, S.; Li, Z. A Model Study on Raw Material Chemical Composition to Predict Sinter Quality Based on GA-RNN. *Comput. Intell. Neurosci.* **2022**, *2022*, 3343427. [[CrossRef](#)] [[PubMed](#)]
9. Song, L.; Jianguang, L.; Fumin, L.; Qing, L. Research on prediction model of basic sintering characteristics of mixed iron ore and sinter quality. In Proceedings of the International Conference on Management, Computer and Education Informatization (MCEI), Shenyang, China, 25–27 September 2015; pp. 187–190.
10. Majumder, A.; Biswas, C.; Dhar, S.; Das, G. Optimization of Process Parameters to Achieve Better Mechanical Properties and Higher Productivity in Sinter Plant. *MGMI Trans.* **2021**, *116*, 48–59.
11. Kinnunen, K.; Laitinen, P. Modeling of Reduction Degradation of Iron Ore Sinter by Feed-Forward Neural Networks. *IFAC Proc. Vol.* **2004**, *37*, 97–102. [[CrossRef](#)]
12. Yuan, Z.; Wang, B.; Liang, K.; Liu, Q.; Zhang, L. Application of deep belief network in prediction of secondary chemical components of sinter. In Proceedings of the 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, China, 31 May–2 June 2018; pp. 2746–2751.
13. Westphal, E.; Seitz, H. A machine learning method for defect detection and visualization in selective laser sintering based on convolutional neural networks. *Addit. Manuf.* **2021**, *41*, 101965. [[CrossRef](#)]
14. Xiao, L.; Lu, M.; Huang, H. Detection of powder bed defects in selective laser sintering using convolutional neural network. *Int. J. Adv. Manuf. Technol.* **2020**, *107*, 2485–2496. [[CrossRef](#)]
15. Xiao, C.; Sun, B.; Wang, Y.; Qiu, L. Foreign Object Detection of Sintering Transport Belt Based on CNN. *IFAC-PapersOnLine* **2021**, *54*, 25–30. [[CrossRef](#)]
16. Ding, Q.; Li, Z.; Zhao, L. FeO Content Classification of Sinter Based on Semi-Supervised Deep Learning. In Proceedings of the 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 22–24 May 2021; pp. 640–644.
17. Frei, M.; Kruijs, F.E. Image-based size analysis of agglomerated and partially sintered particles via convolutional neural networks. *Powder Technol.* **2020**, *360*, 324–336. [[CrossRef](#)]
18. Yu, C.; Han, R.; Song, M.; Liu, C.; Chang, C.I. A simplified 2D-3D CNN architecture for hyperspectral image classification based on spatial-spectral fusion. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 2485–2501. [[CrossRef](#)]
19. Le, N.Q.K.; Ho, Q.T. Deep transformers and convolutional neural network in identifying DNA N6-methyladenine sites in cross-species genomes. *Methods* **2021**, *204*, 199–206. [[CrossRef](#)]
20. Le, N.Q.K.; Nguyen, B.P. Prediction of FMN binding sites in electron transport chains based on 2-D CNN and PSSM Profiles. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2019**, *18*, 2189–2197. [[CrossRef](#)] [[PubMed](#)]
21. Aziz, A.; Sohail, A.; Fahad, L.; Burhan, M.; Wahab, N.; Khan, A. Channel boosted convolutional neural network for classification of mitotic nuclei using histopathological images. In Proceedings of the 2020 17th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 14–18 January 2020; pp. 277–284.
22. Gao, Y.; Li, J.; Liu, C.; Tang, H. Study on Optimizing the Sintering Proportioning by Experiment. In Proceedings of the 8th Pacific Rim International Congress on Advanced Materials and Processing, Waikoloa, HI, USA, 4–9 August 2013; pp. 2221–2229.
23. Shen, X.; Chen, L.; Xia, S.; Xie, Z.; Qin, X. Burdening proportion and new energy-saving technologies analysis and optimization for iron and steel production system. *J. Clean. Prod.* **2018**, *172*, 2153–2166. [[CrossRef](#)]
24. Wu, S.L.; Oliveira, D.; Dai, Y.M.; Xu, J. Ore-blending optimization model for sintering process based on characteristics of iron ores. *Int. J. Miner. Metall. Mater.* **2012**, *19*, 217–224. [[CrossRef](#)]
25. Liu, C.; Xie, Z.; Sun, F.; Chen, L. Optimization for sintering proportioning based on energy value. *Appl. Therm. Eng.* **2016**, *103*, 1087–1094. [[CrossRef](#)]
26. He, Y. Selection of Optimal Ingredients for Out-of-furnace Smelting Based on Computer Application. *J. Phys.* **2020**, *1533*, 022006. [[CrossRef](#)]
27. Liu, S.; Liu, X.; Lyu, Q.; Li, F. Comprehensive system based on a DNN and LSTM for predicting sinter composition. *Appl. Soft Comput.* **2020**, *95*, 106574. [[CrossRef](#)]

28. Wu, M.; Chen, X.; Cao, W.; She, J.; Wang, C. An intelligent integrated optimization system for the proportioning of iron ore in a sintering process. *J. Process Control* **2014**, *24*, 182–202. [[CrossRef](#)]
29. Wang, J.; Ma, Y.; Qiao, F.; Zhai, X.; Liu, J. Energy-Aware Cascade optimization for Proportioning in the Sintering Process Using Improved Immune-Simulated Annealing Algorithm. In Proceedings of the 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy, 6–9 October 2019; pp. 397–402.
30. Wu, T.; Liu, Y.; Tang, W.; Li, X.; Yu, Y. Constraint genetic algorithm and its application in sintering proportioning. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Singapore, 28–30 July 2017; Volume 231, p. 012022.
31. Sun, L.; Wu, X.; Wang, C.; Shi, Z. Research on the optimization algorithm of sintering ingredients. In Proceedings of the IOP Conference Series: Materials Science and Engineering, Nanjing, China, 23–26 May 2018; Volume 382, p. 022089.
32. Chen, J.; Zuo, H.; Xue, Q.; Wang, J. A review of burden distribution models of blast furnace. *Powder Technol.* **2021**, *398*, 117055. [[CrossRef](#)]
33. Deng, L.; Yu, D. Deep Learning—Methods and Applications. *Found. Trends Signal Process.* **2014**, *7*, 197–387. [[CrossRef](#)]
34. Al Arif, S.; Knapp, K.; Slabaugh, G. Spnet: Shape prediction using a fully convolutional neural network. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Granada, Spain, 16–18 September 2018; pp. 430–439.
35. Zhang, Y.; Davison, B.D. Shapenet: Age-focused landmark shape prediction with regressive cnn. In Proceedings of the 2019 International Conference on Content-Based Multimedia Indexing (CBMI), Dublin, Ireland, 4–6 September 2019; pp. 1–6.
36. Zhang, Y.; Li, Q. A regressive convolution neural network and support vector regression model for electricity consumption forecasting. In Proceedings of the Future of Information and Communication Conference, Francisco, CA, USA, 14–15 March 2019; pp. 33–45.
37. Nakagawa, S.; Schielzeth, H. A general and simple method for obtaining R<sup>2</sup> from generalized linear mixed-effects models. *Methods Ecol. Evol.* **2013**, *4*, 133–142. [[CrossRef](#)]
38. Akshay Prassanna, S.; Harshanand, B.; Srishti, B.; Chaitanya, R.; Kirubakaran Nithiya Soundari, S.; Kumar, V.M.; Varshitha Chennamsetti, V.G.; Maurya, P.K. Crop Value Forecasting using Decision Tree Regressor and Models. *Eur. J. Mol. Clin. Med.* **2020**, *7*, 3702–3722.
39. Coulston, J.W.; Blinn, C.E.; Thomas, V.A.; Wynne, R.H. Approximating prediction uncertainty for random forest regression models. *Photogramm. Eng. Remote Sens.* **2016**, *82*, 189–197. [[CrossRef](#)]
40. Dorigo, T.; Guglielmini, S.; Kieseler, J.; Layer, L.; Strong, G.C. Deep Regression of Muon Energy with a K-Nearest Neighbor Algorithm. *arXiv* **2022**, arXiv:2203.02841.
41. Borchani, H.; Varando, G.; Bielza, C.; Larranaga, P. A survey on multi-output regression. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2015**, *5*, 216–233. [[CrossRef](#)]
42. Li, Z.; Zhang, D.; Zhao, X.; Wang, F.; Zhang, B.; Ye, D.; Han, J. A temporally smoothed MLP regression scheme for continuous knee/ankle angles estimation by using multi-channel sEMG. *IEEE Access* **2020**, *8*, 47433–47444. [[CrossRef](#)]
43. Xu, S.; An, X.; Qiao, X.; Zhu, L.; Li, L. Multi-output least-squares support vector regression machines. *Pattern Recognit. Lett.* **2013**, *34*, 1078–1084. [[CrossRef](#)]
44. Melki, G.; Cano, A.; Kecman, V.; Ventura, S. Multi-target support vector regression via correlation regressor chains. *Inf. Sci.* **2017**, *415*, 53–69. [[CrossRef](#)]
45. Mathew, F.; Wang, H.; Montgomery, L.; Kildea, J. Natural language processing and machine learning to assist radiation oncology incident learning. *J. Appl. Clin. Med. Phys.* **2021**, *22*, 172–184. [[CrossRef](#)] [[PubMed](#)]
46. Mallick, A.; Dhara, S.; Rath, S. Application of machine learning algorithms for prediction of sinter machine productivity. *Mach. Learn. Appl.* **2021**, *6*, 100186. [[CrossRef](#)]