上海交通大学学位论文

用运行速度更快的神经网络取代 DrBoxOnline.com 上基于云的计算工具

姓　　名：陈柯冶，曹彦卓，张峰雨，邓铄，朱宇宽

学　　号：5200021910008, 520370910021, 520021910702, 520370910120, 520370910142

导　　师：吴继刚

学　　院：密西根学院

专业名称：电子与计算机工程

申请学位层次：学士

2024 年 08 月

A Dissertation Submitted to

Shanghai Jiao Tong University for Bachelor's Degree

REPLACING A CLOUD BASED COMPUTATION TOOL ON DRBOXONLINE.COM WITH FASTER RUNNING NEURAL NETWORK

Author: Keye Chen, Yanzhuo Cao,

Fengyu Zhang, Shuo Deng, Yukuan Zhu

Supervisor: Jigang Wu

UM-SJTU JOINT INSTITUTE

Shanghai Jiao Tong University

Shanghai, P.R. China

August, 2024

上海交通大学
学位论文原创性声明

本人郑重声明：所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本人完全知晓本声明的法律后果由本人承担。

学位论文作者签名：

日期：　　年　　月　　日

上海交通大学
学位论文使用授权书

本人同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。

本学位论文属于：

□**公开论文**

□**内部论文**，保密□ 1 年/□ 2 年/□ 3 年，过保密期后适用本授权书。

□**秘密论文**，保密____ 年（不超过 10 年），过保密期后适用本授权书。

□**机密论文**，保密____ 年（不超过 20 年），过保密期后适用本授权书。

（请在以上方框内选择打"✓"）

学位论文作者签名：　　　　　　　　指导教师签名：

日期：　　年　　月　　日　　　　　日期：　　年　　月　　日

摘　要

本项目旨在解决包装和配送行业中的一个关键挑战，专注于瓦楞纸盒的结构性能。这些盒子因其轻质、可回收和高度可定制性而广受欢迎，但它们在存储和运输过程中容易变形，这不仅可能导致产品损坏，还可能造成收入损失、资源浪费和客户投诉。

目前，行业专家可利用 Dr. Box Calculator Pro 进行有限元分析（FEA），以预测在不同条件下的箱子性能。尽管这种方法有效，但它既耗时又需要大量的计算资源。为了克服这些限制，我们提出了一种创新的解决方案：将深度神经网络（DNN）技术集成到 Dr. Box Calculator Pro 中，以替代部分 FEA 计算过程。

深度神经网络以其多层结构模拟复杂非线性关系的能力而著称，它们能够自动学习特征，提供可扩展性和端到端学习的优势，非常适合处理大规模数据集。

我们的解决方案的目标是实现即时的弯曲强度预测，同时降低计算成本，并有效应对各种盒子类型和条件。我们期望通过这一方法，不仅能够提高预测的准确性，加快分析速度，还能提升包装和物流操作的整体效率。我们面临的主要挑战是开发一个既准确又高效的弯曲强度预测模型，具体目标包括实现高达 90% 以上的预测精度、每次模拟不超过 2 秒的快速分析速率、能够适应特定的盒子配置，以及在实际应用中达到 99.9% 以上的系统可靠性。

预计通过将 DNN 技术集成到现有系统中，我们将能够为包装行业带来革命性的变革，提供更快、更准确、更可扩展的弯曲强度预测。这不仅将降低运营成本，还将通过减少运输过程中的产品损坏来提高客户满意度。本项目的驱动力在于应用前沿的机器学习技术，以解决长期存在的行业问题，并提升包装和物流操作的效率与效果。

关键词：瓦楞纸盒，深度神经网络，有限元分析，弯曲强度预测

ABSTRACT

This project tackles a critical issue in the packaging and delivery industry, focusing on structural behaviors of corrugated paper boxes. Although these boxes are widely used because of their lightweight, recyclable, and customizable properties, they are prone to buckling during storage and shipment, resulting in product damage, lost revenue, wasted resources, and customer complaints.

Currently, the industry can refer to the Dr. Box Calculator Pro, which employs Finite Element Analysis (FEA) to predict box performance under various conditions. While effective, FEA is time-consuming and demands significant computational resources. To address these limitations, we propose integrating Deep Neural Networks (DNN) into the Dr. Box Calculator Pro to replace part of the FEA calculation. DNN's ability to model complex non-linear relationships through multiple layers offers advantages such as automatic feature learning, scalability, and end-to-end learning, making them well-suited for big data applications.

Our solution aims to provide instant predictions of buckling strength, reduce computational costs, and efficiently handle diverse box types and conditions. The expected outcome is enhanced prediction accuracy, faster analysis, and overall improved efficiency in packaging and logistics operations. The primary challenge is achieving accurate and efficient predictions of buckling strength, with specifications including high predictive accuracy ($\geq$ 90%), fast analysis rate ($\leq$ 2 seconds per simulation), the ability to handle a specific box configurations, and system reliability above 99.9% in practice.

The integration of DNNs is anticipated to revolutionize the packaging industry by offering faster, more accurate, and scalable predictions of buckling strength. This advancement will reduce operational costs and enhance customer satisfaction by minimizing product damage during transit. The motivation behind this project is to apply modern machine learning techniques to solve long-standing industry problems, improving the efficiency and effectiveness of packaging and logistics operations.

Key words: Corrugated Paper Box, DNN, FEA, Buckling Strength Prediction

Contents

Chapter 1 Introduction

1.1 Background

Our project addresses critical challenges in the packaging and logistics industry, particularly concerning corrugated paper boxes. These boxes are extensively used due to their lightweight, recyclable, and customizable properties. However, they are prone to buckling during storage and shipment, leading to product damage, revenue loss, wasted resources, and customer complaints.

One of the current solution to this problem is the Dr. Box Calculator Pro, a website that provides users with an interface to predict the performance of corrugated boxes using Finite Element Analysis (FEA). As shown in Figure 1–1 FEA is a numerical method for predicting how a product reacts to real-world forces, vibrations, heat, and other physical effects by dividing a complex problem into smaller, simpler parts (finite elements) and solving them. Despite its effectiveness, FEA is time-consuming and demands extensive computational resources.

Our proposed solution involves replacing FEA with Deep Neural Networks (DNN) in the Dr. Box Calculator Pro[1]. DNNs are artificial neural networks with multiple layers between the input and output layers, allowing them to model complex non-linear relationships. DNNs offer several advantages, including automatic feature learning from raw data, scalability to handle large-scale data and complex models, and the capability for end-to-end learning, which simplifies the prediction pipeline.

Our expected outcomes are improved accuracy, speed and lower inference cost. By incorporating DNNs into the Dr. Box Calculator Pro, we can achieve instant predictions of buckling strength, significantly reduce computational costs, and expedite analysis. Additionally, DNNs can handle diverse box types and conditions, enhancing the efficiency and effectiveness of packaging and logistics operations.

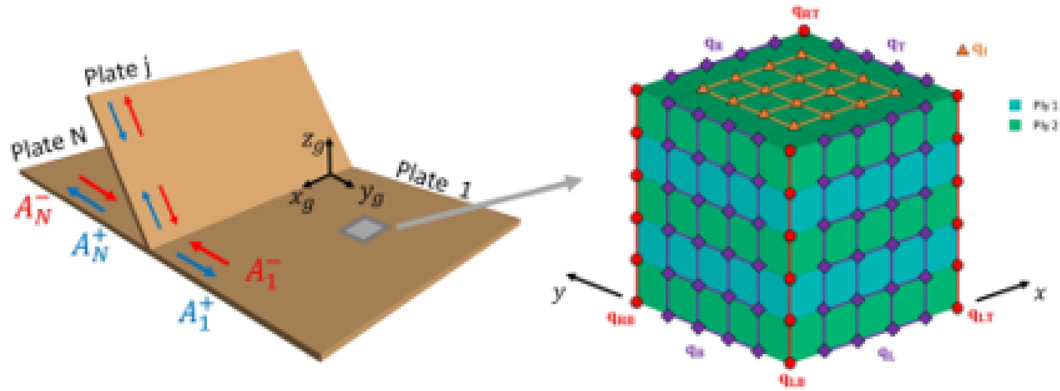


Figure 1–1 FEA for Corrugated Paper Box

1.2 Literature Review

1.2.1 Traditional Finite Element Analysis (FEA) Approaches

Finite Element Analysis (FEA) has been extensively used to predict the buckling strength of corrugated cardboard boxes, offering a robust and accurate method for such analyses. For instance, as shown in Figure 1–2, Feher et al.[2] analyzed the compressive strength of corrugated paperboard packages with different cutout rates using FEA. Their study, which employed a homogenized linear elastic orthotropic material model with Hill plasticity and modeled only one-eighth of the box to simplify analysis, demonstrated high accuracy in predicting compression force from Box Compression Tests (BCT) despite a slight decrease in accuracy at higher cutout rates.

Traditional methods for analyzing the strength and failure of corrugated boxes include experiments, analytical models, and computational software. However, these methods require extensive experimentation and are often limited to specific box types and loading conditions. For example, analytical models like the McKee equation are limited to vertically loaded regular slotted containers (RSC).

1.2.2 Integration of Artificial Neural Networks (ANN) with FEA

A significant advancement in predictive modeling involves integrating Artificial Neural Networks (ANN) with FEA. This approach has been explored to improve prediction accuracy and computational efficiency, as evidenced by study of Hajializadeh[3] on predicting residual

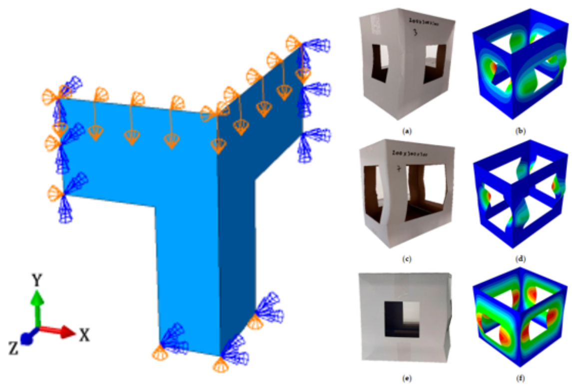


Figure 1–2 Numerical Analysis of FEA by Feher et al.[2]

stresses in AISI 304 L parts created through the Direct Metal Deposition (DMD) process. Their innovative model integrated ANN with FEA, focusing on the thermal and mechanical response of materials. The ANN was trained using datasets generated from FE simulations, including spatial coordinates and temperature history, to predict residual stresses. As also shown in Figure 1–3, this method not only enhanced prediction accuracy but also significantly reduced computational time.

The novel aspect of integrating ANN with FEA lies in its ability to handle complex data relationships more effectively than traditional methods. This integration allows for quicker evaluations and can be adapted to different box types and materials, providing a flexible and scalable solution for packaging analysis.

1.2.3 Comparison and Novelty of the Proposed Design

In comparison to the aforementioned studies, our proposed design integrates the strengths of both traditional FEA and ANN to provide a comprehensive and efficient analysis tool for corrugated cardboard boxes. Unlike the study by Feher[2], which solely relied on FEA, and the work of Hajializadeh[3], which focused on metal materials, our design applies ANN-enhanced FEA specifically to corrugated cardboard boxes. This approach addresses the limitations of traditional methods by improving prediction accuracy and reducing computational time, making it suitable for a wider range of packaging solutions.

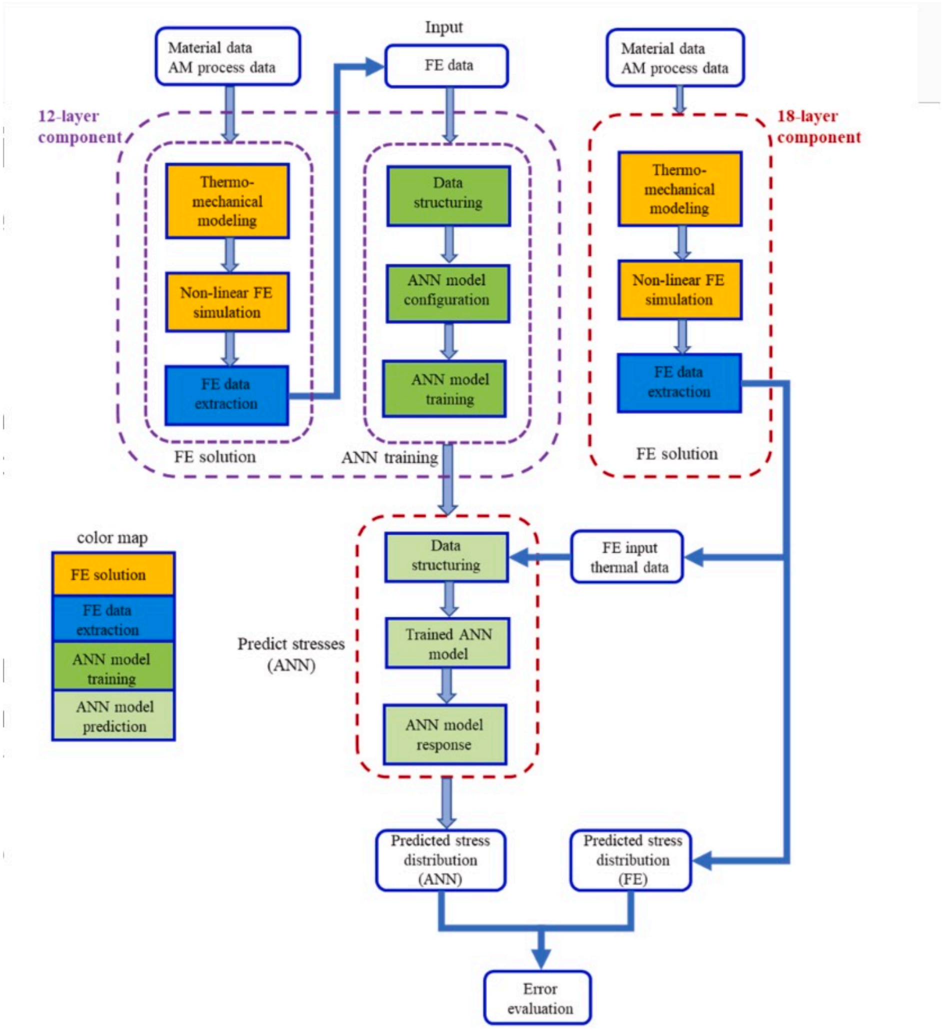


Figure 1–3 Training Process for DMD[3]

The novelty of our design lies in its ability to predict box performance based on paper properties for various box types and designs, including special features like hand holds and ventilation holes. Moreover, in contrast, our model is further designed for multi-task labeling, which includes nodal Deformation x, nodal Deformation y, nodal Deformation z, nodal Failure, displacement and force. Each dimension has more 128 outputs to generate. Like Figure 1–4, it also shows a broader application in the physical field.

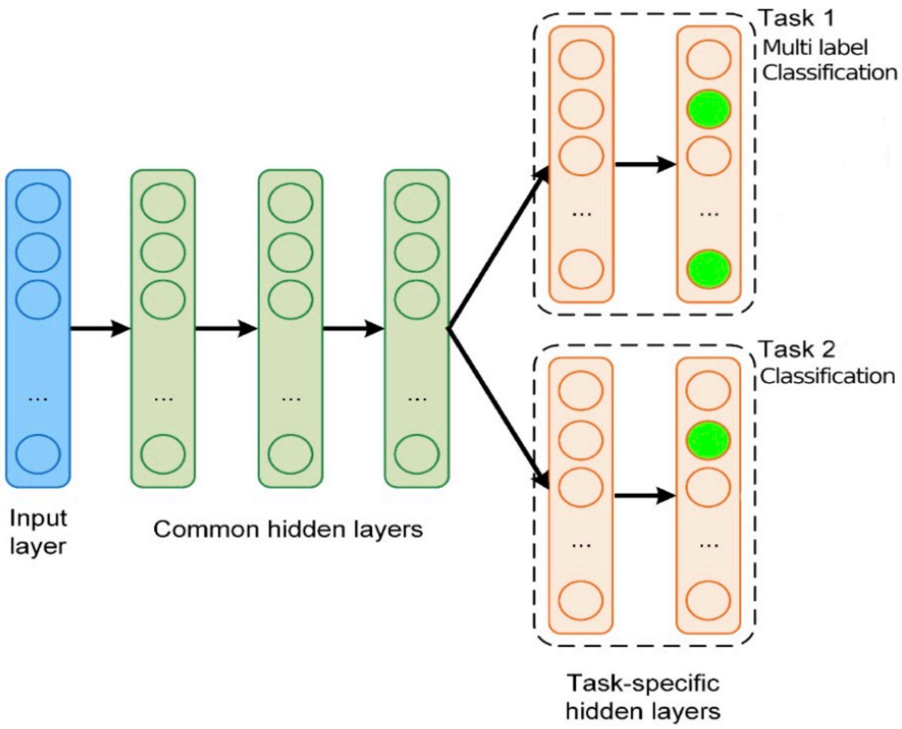


Figure 1–4 Multi-task Predictions[3]

This hybrid approach combines the accuracy of FEA with the efficiency of ANN, enabling non-specialized personnel to conduct robust and cost-effective packaging analyses.

1.3 Significance

Our project holds great significance for the packaging and logistics industry, as it could improve the way packaging performance is predicted and analyzed. By integrating DNN with FEA, we are not only advancing the state of the art in packaging technology but also setting a precedent for solving complex problems across various domains.

Economically, our solution is poised to deliver substantial benefits by reducing operational costs and enhancing customer satisfaction, thereby bolstering the competitive edge of businesses.

The environmental sustainability aspect of our project is equally important, as it aligns with the global push towards eco-friendly solutions. By minimizing product damage during transit, we contribute to reducing waste and promoting the circular use of resources.

The practical applicability of our solution is of paramount importance, as it offers immediate predictions at a fraction of the computational cost, a feature that is invaluable in industries where swift decision-making is crucial.

Furthermore, our project underscores the power of data-driven decision-making, empowering businesses to make more accurate choices based on the automated feature learning capabilities of DNNs. This, in turn, optimizes product design and logistics strategies.

Additionally, the societal impact of our project should not be underestimated. By mitigating the inconvenience and dissatisfaction of consumers caused by damaged goods in transit, we aim to elevate the overall satisfaction within society.

In conclusion, the significance of our project extends beyond mere technological innovation, touching upon economic, environmental, and social aspects, making it a holistic and impactful endeavor.

Chapter 2 Design Specifications

2.1 Customer Requirements

At the beginning, the customer requirements for us were only relatively high accuracy. After many offline discussions, we came up with the following detailed customer requirements.

- **Good Simulation:** High-quality predictions that accurately reflect real-world outcomes.
- **Very Fast:** The system should deliver results rapidly to enable efficient decision-making.
- **Can Accept Various Inputs:** The system must be capable of handling diverse input variables to accommodate different user needs.
- **Very Reliable:** The system needs to operate with high reliability and minimal downtime.
- **User-Friendly :** The interface should be easy to use for all user levels, ensuring a low learning curve.
- **Inexpensive:** The solution should be cost-effective for the user.

2.2 Process to Determine Engineering Targets

After confirming the detailed customer requirements, we held several internal seminars to determine the engineering specifications and the granularity.

- **Good Simulation (Accuracy, Feature Learning, End-to-End Learning)**
  - *Requirement Translation*: Convert the need for ”high-quality predictions” into a specific accuracy metric.
  - *Specification*: Achieve at least 90% accuracy in simulation predictions.
  - *Metric*: Use MSE loss, cosine cimilarity loss and length difference to measure.
  - *Process*: Involve feature learning capabilities in the DNN to automatically discern and utilize the most impactful features from input data, enhancing prediction accuracy through deep learning techniques.

- **Very Fast (Speed, System Optimization)**
  - *Requirement Translation*: Define what ”deliver results rapidly” means in a measurable time frame.
  - *Specification*: Reduce simulation time to under 2 seconds per run.
  - *Metric*: Implement benchmarking tests to track and verify the response times across various operations within the system.
  - *Process*: Optimize the computational efficiency of the DNN model by refining algorithmic approaches and utilizing efficient computing resources.
- **Can Accept Various Inputs (Scalability)**
  - *Requirement Translation*: Specify the types and variety of inputs the system should handle.
  - *Specification*: Support at least 3 distinct types of input variables.
  - *Metric*: Test the system’s ability to process and accurately analyze different data types and sources.
  - *Process*: Design the system architecture to be modular, allowing easy integration of new input variables without significant modifications.
- **Very Reliable (Server Reliability, System Optimization)**
  - *Requirement Translation*: Translate the need for high reliability into specific uptime and redundancy requirements.
  - *Specification*: Achieve 99.9% system uptime.
  - *Metric*: Use server monitoring tools to track uptime and automatically detect and resolve potential downtimes.
  - *Process*: Implement fault tolerance mechanisms such as redundancy and regular maintenance checks to ensure consistent performance.
- **User-Friendly (User Interface Design)**
  - *Requirement Translation*: Define what makes the interface ”easy to use”.
  - *Specification*: Achieve a user satisfaction rating of at least 80% on usability.
  - *Metric*: Conduct user testing sessions and gather feedback to assess the intuitiveness of the interface.
  - *Process*: Apply principles of user-centered design during the development phase to ensure the interface is intuitive and accessible to all user levels.

- **Inexpensive (Cost)**
  - *Requirement Translation*: Quantify what "cost-effective" means in comparison to competing solutions.
  - *Specification*: Keep the cost per simulation below a competitive benchmark, say 20% less than similar services.
  - *Metric*: Regularly review cost structures and market prices to ensure competitiveness.
  - *Process*: Streamline operations and utilize cost-effective technologies and methodologies to reduce overheads and pass these savings on to users.

2.3 Quality Function Deployment (QFD) and Specific Engineering Requirements

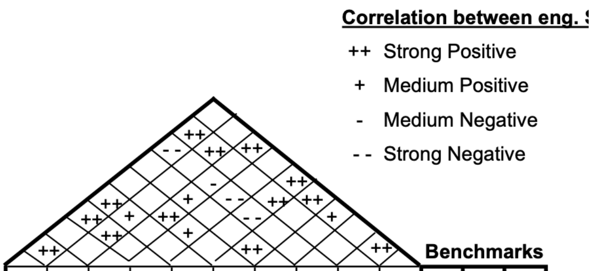
2.3.1 Critical-to-Quality Engineering Specifications

In this section, we delve into the critical-to-quality engineering specifications derived from our QFD analysis. The QFD diagram, as depicted in Table 2–1, illustrates the correlation between customer requirements and engineering specifications, which are pivotal in ensuring product quality and meeting customer expectations.

The weights assigned to each engineering specification are indicative of their importance in the overall product development process. For instance, 'Good simulation' and 'Very fast' have been assigned high weights, reflecting their critical nature in satisfying customer needs.

For "Good simulation": This specification has a strong positive correlation with customer requirements, indicating that a high-quality simulation is essential for meeting customer expectations. The normalized weight of 0.13 underscores its significance. For "Very fast": Speed is another critical factor, with a normalized weight of 0.08, suggesting that the system's responsiveness is crucial for user satisfaction. For "Accept various inputs": This specification is also highly valued, with a normalized weight of 0.07, indicating the importance of flexibility in handling diverse input types. For "Very reliable": Reliability is a key concern, reflected in the normalized weight of 0.09, highlighting the need for a dependable system. For "User friendly": With a normalized weight of 0.10, user-friendliness is a significant factor in ensuring ease of use and overall customer satisfaction. For "Inexpensive":

Table 2–1 QFD Diagram



	Weight (1-10)	DNN Integration	Feature Learning	Scalability	End-to-End Learning	System Optimization	Server Reliability	User Interface Design	Cost	Accuracy	Speed	Benchmarks: competitor A	competitor B	
Good simulation	10	10	8		5	10	7			10		9	8	
Very fast	8	8	7	7	7	10		7	7		10	5	4	
Accept various inputs	8	7	6	10	8	7		7		7		10	7	
Very reliable	7	5			4	8	10		7	5	6	9	7	
User friendly	5				3	6	6	10				5	4	
Inexpensive	7	7		5		10	8	8	10	6	7	5	4	
Measurement Unit		MSE	Score	GB	h	LOP	%	Score	¥	%	s			
Target Value							99	20		90	300			
Total		304	184	171	213	392	226	218	175	233	171			
Normalized		0.13	0.08	0.07	0.09	0.17	0.10	0.10	0.08	0.10	0.07			

Correlation between customer reqs and eng. Specs
9 = Strong Relationship
3 = Medium Relationship
1 = Small Relationship
(blank) = Not Related

Cost-effectiveness, represented by a normalized weight of 0.07, is another important consideration, balancing quality with affordability.

2.3.2 Specific Engineering Requirements

Since our customers want to ensure all the requirements quantitatively and all the measuring method clearly, after a detailed discussion with the customer, we made another table of specific engineering requirements as below as Table 2–2.

Table 2–2 Table of Specific Engineering Requirements

Requirements	Specifications	Target	Measurement
Accuracy	Predictive accuracy within 90%	≥ 90% accuracy	Statistical analysis of prediction vs actual
Speed	Analysis time per simulation	≤ 5 minutes	Time tracking software
Customization	Support diverse conditions	Handle 20+ variables	System configuration checks
Reliability	System operational uptime	99.9% uptime	Server Monitoring tools

Chapter 3 Concept Generation

3.1 Data Augmentation

From the perspective of customization, some sub-function includes stable prediction for boxes with different scales and the ability to capture environmental symmetry. We came up with several ideas to meet these requirements.

- Oversampling: Different sampling probability can be assigned to the data bins respectively based on their frequency to balance the distribution of training data.
- Customized Loss Function: We can assign larger weight to sparsely distributed data points when calculating the loss function. This can make minority data more important during back propagation.
- Data Augmentation: By generating new data with x-y spatial symmetry from the original ones, we allow the model to learn spatial symmetry conditions while relieving the imbalance by expanding the training data.

Finally, we proposed augmenting data with spatial symmetry, which would allow the model to learn symmetry conditions while relieving the imbalance by expanding the training dataset.

3.2 Activation Function

To achieve system reliability above 99.9%, our solution need to ensure stable converging performance during repeated model training sessions, acceptable variance of (MSE) loss across multiple validation sets, and maintaining high reliability when handling extremely large or small inputs. Since the only non-linearly is induced by the activation functions, we basically discuss the choices of common activation functions.

- Sigmoid

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

- Hyperbolic Tangent

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

- Relu Function

$$\mathrm{ReLU}(x) = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases}$$

- Leaky ReLU

$$\mathrm{LeakyReLU}(x) = \max(\alpha x, x)$$

- Gaussian Error Linear Unit

$$\mathrm{GELU}(x) = \frac{x}{2} \cdot (1 + \tanh(\sqrt{2/\pi} \cdot (x + 0.044715 \cdot x^3)))$$

After intense tests on the model, we founded that relu out-performs in regular regression model while sigmoid out-performing in probability models.

3.3 Dimension Reduction

As we deal with high-dimensional data outputs, such as 128-dimensional vectors representing various deformation parameters, dimension reduction becomes essential. By dimension reduction such as RCA and t-sne, we may be able to effectively reduce the dimensionality from 128 to a more manageable 1 to 20 dimensions. This not only aims to simplify the data, making it easier to analyze and interpret, but also try to preserve the global structure of the data. This is critical for understanding overarching patterns and trends in the data, which are fundamental for accurate modeling and prediction.

3.4 Optimizer Selection

The choice of an optimizer is integral to the training of neural networks, particularly in terms of convergence speed and the model’ s performance on new, unseen data. We have optimizer options of SGD, RMSprop as well as AdamW, with their adaptive learning rate and weight decay strategy, ensures that our model not only learns efficiently but also generalizes better to other data. This is especially important in a project where stability and reliability in predictions are paramount. By optimizing these parameters, we can have more stable and effective learning, which is critical in achieving high accuracy and performance in our neural network applications.

3.5 Overall Generated Concepts

As shown in Table 3–1, it is our group's overall generated concepts.

Table 3–1 Overall Generated Concepts

Function	Options	Advantages	Disadvantages
Data Augmentation	• Oversampling • Customized Loss Function • Data Augmentation with x-y spatial symmetry	• Balances training data distribution • Increases importance of minority data • Enhances model's understanding of spatial symmetry	• May introduce bias • Could overfit on minority samples • May not generalize well beyond trained symmetries
Activation Function	• More complex models • Smoother functions like ReLU • Anomaly detection with exception handling	• Prevents extreme values • Stable performance • Warns users of anomalies	• May limit model capacity • Simple functions may not capture complex patterns • Additional computational overhead
Dimension Reduction	• PCA • t-SNE	• Simplifies high-dimensional data • Preserves global data structure	• t-SNE may not scale well with larger datasets • PCA assumes linearity
Optimizer Selection	• SGD • RMSprop • AdamW	• Adapts learning rate • Efficient learning and generalization	• SGD may converge slowly • RMSprop might not perform well on non-stationary problems • AdamW could lead to overfitting in some cases

Chapter 4 Concept Selection

In the context of conceptual selection methodology, we will first introduce our concept and aim with respective to engineering specifications to qualitatively and quantitatively analyze our chosen subject and its advantages and disadvantages. This will enable our neural network more accurate and robust, handling limited size inputs to output high dimension result.

4.1 Data Augmentation

We evaluated several strategies to enhance the training dataset for a neural network model predicting the structural behavior of corrugated paper boxes. The decision was made to select ”Data Augmentation with x-y spatial symmetry.” The following outlines the reasons for not selecting the other considered options: ”Oversampling” and ”Customized Loss Function.”

- **Potential for Overfitting:** Replicating minority class samples excessively can lead the model to overfit these scenarios, which can degrade its performance on new, unseen data.
- **Skewed Data Distribution:** Oversampling does not introduce new information; it merely repeats what is already present, which might not be sufficient for learning complex, varied real-world scenarios.
- **Complexity in Implementation:** Designing and tuning a loss function that weights minority data more significantly is complex and can lead to an imbalance in learning priorities.
- **Balancing Act:** Improper weighting can cause the model to bias towards less frequent patterns at the expense of overlooking common, critical patterns.

As for the advantages of choosing data augmentation with x-y spatial symmetry, which is also shown in Concept Selection Matrix. This method introduces new, practical data variations, enhancing the model’s ability to generalize from training to real-world applications. Additionally, augmenting data by simulating physical changes such as orientation relative to forces enriches the model’s understanding of structural behaviors under varied conditions.

This approach allows the model to develop a nuanced understanding of box behaviors, leading to improved accuracy and performance on diverse datasets.

4.2 Activation Function

This section describes the rationale behind the selection of the activation function for a neural network model designed to predict the structural behaviors of corrugated paper boxes. The decision was made to employ ”ReLU (Rectified Linear Unit)” enhanced with ”Anomaly Detection and Exception Handling.” Compared to more complex activation functions or models, the decision has following advantages:

- **Simplicity and Efficiency:** ReLU is preferred for its simplicity and efficiency in training deep neural networks. It accelerates the convergence of stochastic gradient descent compared to sigmoid or tanh functions due to its linear, non-saturating form.
- **Non-linearity:** While ReLU introduces non-linearity into the model, it maintains a range that allows for effective gradient propagation, crucial for learning complex patterns without the vanishing gradient problem.
- **Enhancing Robustness:** To safeguard the model against potential anomalies during training and inference, anomaly detection mechanisms will monitor activation outputs for unexpected behaviors, such as excessively high or zero values indicative of dying ReLU problems.
- **Exception Handling:** In cases where anomalies are detected, exception handling procedures are implemented to adjust computations or revert to safer states, ensuring model stability and reliability.

However, more complex activation functions like parametric ReLU or ELU (Exponential Linear Unit) could increase the model’s capacity but also its susceptibility to overfitting and higher computational cost. For the project’s scope, which requires balancing performance with computational efficiency, ReLU provides an optimal compromise without unnecessary complexity.

Thus, the choice of ReLU, enhanced with anomaly detection and exception handling, aligns with the project’s goals to develop a robust, efficient, and reliable model. This approach not only simplifies the learning process but also ensures that the network remains stable under various operating conditions, crucial for the predictive analysis of box structures.

4.3 Diemnsion Reduction

Our model outputs features like x, y, z deformation, force, nodal failure, and displacement, each as 128-dimensional vectors containing deformation information of various nodes in a cardboard box. To manage and analyze this high-dimensional data, we need to reduce its dimension from 128 to between 1 and 20 dimensions. We have chosen Principal Component Analysis (PCA) over t-SNE for this purpose due to several reasons:

1. **Interpretability:** PCA simplifies the high-dimensional data by identifying the principal components that account for the most variance, making the outputs easier to understand.
2. **Linear Dimensionality Reduction:** PCA is suitable for our data assuming linear relationships among the deformation parameters. It effectively captures these linear or near-linear relationships.
3. **Preservation of Global Structure:** Unlike t-SNE, which focuses on local relationships, PCA maintains the global structure of the data, a critical aspect in understanding overall deformation patterns and trends.

4.4 Optimizer Selection

Moreover, in neural network training, selecting the right optimizer is crucial for model performance and convergence speed. AdamW is a popular optimizer that combines the benefits of Adam's approach with an improved weight decay strategy[4], enhancing generalization in some scenarios. Here we discuss the reasons for choosing AdamW over SGD or RMSprop from a mathematical perspective. The update rule for AdamW can be expressed as:

$$\theta_{t+1} = \theta_t - \eta \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} - \eta \cdot \lambda \cdot \theta_t$$

where θ_t represents the parameters at step t, η is the learning rate, $\hat{m}_t$ and $\hat{v}_t$ are bias-corrected estimates of the first and second moments, ϵ is a small constant for numerical stability, and λ is the weight decay coefficient.

- **Adaptive Learning Rate:**
  - **SGD** (Stochastic Gradient Descent) uses the basic update rule $\theta_{t+1} = \theta_t - \eta \cdot g_t$, where g_t is the gradient. It lacks an adaptive learning rate mechanism, potentially leading to imbalanced training dynamics across different parameter dimensions.

  - **RMSprop** incorporates an exponential moving average of squared gradients to adjust learning rates per dimension, reducing oscillations during training. Its update rule is $\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{v_t+\epsilon}} \cdot g_t$, where v_t is the moving average of squared gradients.

AdamW combines advantages from both by using estimates of the first and second moments to adaptively adjust the learning rate for each parameter, promoting more stable and effective optimization, and offering superior convergence speeds and generalization compared to SGD and RMSprop.

4.5 Concept Selection Matrix

Based on our generated concepts, we present three different configurations for the neural network model designed to enhance the predictive analysis of structural behaviors in corrugated paper boxes. Each model incorporates a unique combination of data augmentation techniques, activation functions, dimension reduction methods, and optimizers to optimize performance for specific scenarios:

- **Model A**:
  - Data Augmentation with x-y spatial symmetry
  - Smoother Functions like ReLU and Anomaly Detection with Exception Handling
  - PCA Dimension Reduction Method
  - AdamW Optimizer
- **Model B**:
  - Oversampling
  - More Complex Model
  - t-SNE Dimension Reduction Method
  - SGD Optimizer
- **Model C**:
  - Customized Loss Function
  - Smoother Functions like ReLU
  - PCA Dimension Reduction Method
  - RMSprop Optimizer

After our simulation, we decided to use Model A (Data Augmentation with x-y spatial symmetry, Smoother Functions like ReLU and Anomaly Detection with Exception Handling, PCA Dimension Reduction Method, AdamW Optimizer) as our selected models as well as concepts since Model A obtained the highest score according to Table 4–1.

Table 4–1 Concept Diagram

Design criterion	Weight Factor	Unit	Model A			Model B			Model C		
			Value	Score	Rating	Value	Score	Rating	Value	Score	Rating
high prediction accuracy	0.36	%	95	9	3.24	90	7	2.52	92	8	2.88
Feature Learning	0.07	Score	Good	8	0.56	Good	8	0.56	Good	7	0.49
Scalability	0.05	Score	Excellent	8	0.4	Good	6	0.3	Good	6	0.3
System Optimization	0.06	Score	Excellent	9	0.54	Good	6	0.36	Good	5	0.3
Server Reliability	0.12	%	99.9	9	1.08	99	9	1.08	90	6	0.72
User Interface Design	0.02	Score	Fair	6	0.12	Fair	6	0.12	Fair	6	0.12
Cost	0.12	¥	20	9	1.08	40	7	0.84	30	5	0.6
Speed	0.2	s	260	8	1.6	200	9	1.8	400	6	1.2
			8.62			7.58			6.61		

Chapter 5 Final Design

According to our previous selection process, we decided to use a model with data augmentation with x-y spatial symmetry, smoother functions like ReLU and anomaly detection with exception handling, PCA dimension reduction method and AdamW optimizer. The detailed layout drawing is shown in Figure 5–1.

5.1 Engineer Design Analysis

5.1.1 General Design Principles

Based on the chosen concepts, we designed models with various structure and test them intensely to tune the weights. During the process, we followed these principles and methods for engineer design:

- Structure: Our models' structure should be adapted to inputs and outputs. For example, we had an output dimension of either 128 representing the number of node or 10 representing the number of principal components when predicting nodal behaviors. Also, the structure should reflect real-world mechanisms.
- Complexity: For model complexity trade-off, adding complexity can better fit the relationship, while reducing complexity can prevent overfitting and boost inference speed. We should adjust the parameters according to the training loss and evaluation loss curves.
- Data: Cross validation is a common practice to evaluate the performance of a model. We randomly splited 20% of the data for cross-validation. Since the data is limited, we also performed data augmentation with spatial symmetry to improve the performance.

5.1.2 Model Design Development

We first tried models without PCA to explore how the models could capture the relationships between node outputs without dimensional reduction.

- Basic MLP model without PCA: We started with the most basic full connection structure, with the basic module of linear layer, relu function, batch normalization and dropout repeated for 3 times. The basic model only had a parameter number of about 270 thousand. It could be seen that the loss on the training set remained high during the training process, indicating that the parameters struggled to represent the relationship between the inputs and outputs. A more complex model was needed.

- Independent layer with self-attention: To capture independence while preventing the model complexity from intense increase, we use a single independent linear layer, followed by self-attention modules simulating the dependency between 128 nodes. The structure of the model accords with the essence of FEA: first assign the nodes and then calculate the interaction. We reduced the parameter num to around 800 thousand, while improving the accuracy to 840%. Nevertheless, we still noticed slight fluctuation at the end of the evaluation loss, which indicated that there might be overfitting.
- Independent layer with common MLP part: To further reduce the complexity, we reserve the independent layer part, while replacing the self-attention part with a common MLP module to capture the dependency. The parameter number was reduced to 430 thousand and the accuracy increased to 0.863
- MLP model with PCA: With PCA we successfully obtained a linear transformation to convert the 128 dimensions into 10 principal components while remaining a variance explanation ratio more than 99.8%. After experiment, we found that MLP model, along with batch normalization and dropout was capable enough for prediction. To reduce the significant scale difference between the value of each node, we normalize the matrix with column standardization before PCA, and use whiten algorithms to ensure balanced variance among the principal components.

5.2 Design Description

The chosen concept integrates a Deep Neural Network (DNN) within the Dr. Box Calculator Pro system, aimed at enhancing the prediction of buckling strength for corrugated paper boxes. This new approach replaces traditional Finite Element Analysis (FEA) with advanced neural network technologies that accelerate processing times and improve accuracy.

System Overview and Operation (Refer to Figure 5–1)

User Interaction Interface:

- **Input Data:** Users input box specifications, including dimensions, material type, and expected load conditions. This interface is designed for ease of use, ensuring that all necessary data is accurately captured for processing.
- **Result Display:** The system outputs predictions such as maximum load capacity and potential nodal failure points. These results are presented in an accessible format to aid users in decision-making processes.

Training Data:

The DNN is trained on a large dataset comprising historical data that details various box specifications and their performance outcomes. This dataset is vital for training the model to recognize complex relationships and patterns that influence box performance.

Neural Network Model:

- **Hidden Layers:** Specific hidden layers are designated for different prediction tasks—max force and nodal failure—allowing the model to handle diverse aspects of mechanical behavior in corrugated boxes.
- **Optimization:** The network utilizes optimization algorithms to adjust its parameters, aiming to minimize loss functions that directly impact prediction accuracy. Since we have to predict different outcomes with different features as well as different requirements, here we list the two most representative prediction models:
  - **Max Force:** Focuses on minimizing Mean Squared Error (MSE) Loss to refine force prediction.
  - **Nodal Failure:** Targets a reduction in a combined MSE and cosine similarity loss, enhancing the accuracy of failure point predictions.

Implementation Details:

- **Data Processing:** Upon entry, user data is preprocessed to conform to the model’s input requirements, ensuring optimal performance.
- **Prediction Processing:** The DNN processes the input data, applying learned patterns

and behaviors to generate predictions which are then relayed back to the user interface.

Technical Specifications:

- **Back-end Frameworks:** The model operates on a Python-based architecture utilizing TensorFlow, which supports extensive model training and execution.
- **Front-end Development**: The user interface leverages ReactJS to provide a responsive and dynamic experience, facilitating straightforward user interactions.

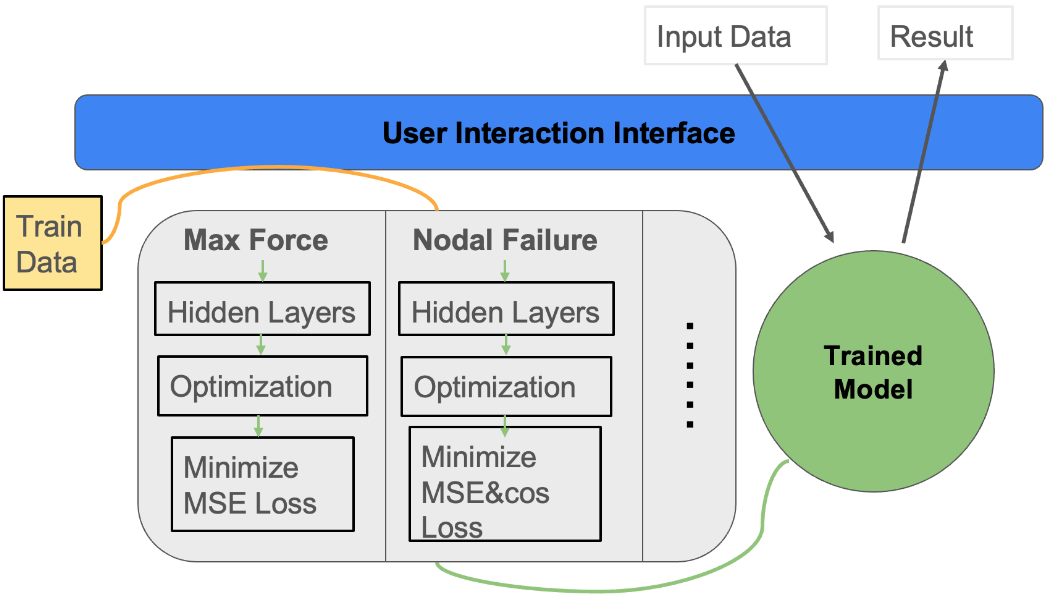


Figure 5–1 Concept Diagram

5.3 Manufacturing Plan

In general, the manufacturing plan should be shown as Figure 5–2, and the detailed plan is described as follows.

5.3.1 Resource Procurement

Hardware Requirements

- **High-Performance Servers:** Acquire high-performance servers with GPU capabilities to handle the computational demands of training and running the DNN models.
- **Storage Solutions:** Ensure ample storage capacity for large datasets and model check-

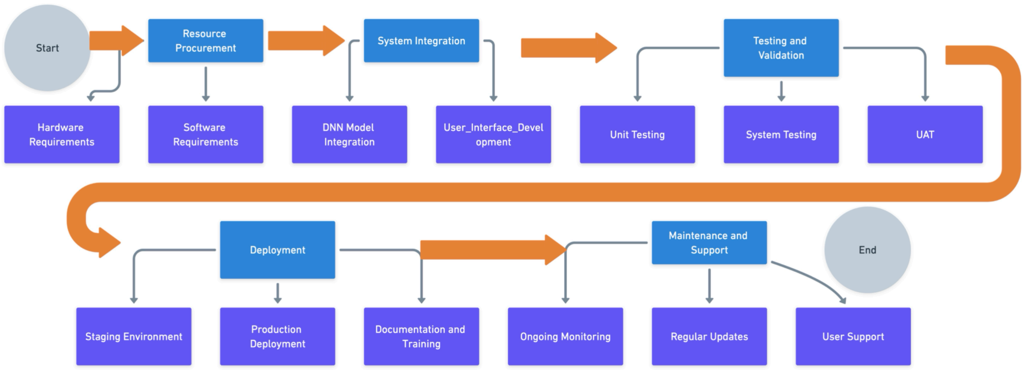


Figure 5–2 Manufacturing Steps

points.

- **Networking Equipment:** High-speed networking equipment to facilitate data transfer and communication between servers.

Software Requirements

- **Development Tools:** Set up development environments with necessary software such as Python, TensorFlow, PyTorch, and other relevant machine learning libraries.
- **Database Management Systems:** Implement robust database management systems for storing and managing training data, model parameters, and simulation results.
- **Version Control:** Use version control systems like Git for managing code and collaborative development.

5.3.2 System Integration

DNN Model Integration

- **Model Training:** Utilize the procured hardware to train the DNN models using prepared and cleaned data. Ensure that the training process includes extensive testing to achieve the desired accuracy and performance.
- **API Development:** Develop APIs to allow the DNN models to interface seamlessly with the Dr. Box Calculator Pro. These APIs will handle input data, run predictions, and return results to the users.

User Interface (UI) Development

- **Front-End Design:** Create a user-friendly interface that allows users to input data and view prediction results. Ensure the interface is intuitive and accessible to users with varying levels of technical expertise.
- **Back-End Integration:** Integrate the back-end systems with the front-end interface. Ensure that the data flow between the UI and the DNN models is smooth and efficient.

5.3.3 Testing and Validation

Unit Testing

- Test individual components of the system, including the DNN models, APIs, and UI elements, to ensure they function correctly in isolation.

System Testing

- Conduct comprehensive testing of the integrated system. Simulate various scenarios to ensure the system performs well under different conditions and inputs.
- Perform stress testing to assess the system's ability to handle high loads and maintain performance.

User Acceptance Testing (UAT)

- Involve end-users in testing the system to gather feedback on its functionality, usability, and performance. Make necessary adjustments based on user feedback to enhance the overall user experience.

5.3.4 Deployment

Staging Environment

- Deploy the system in a staging environment that mirrors the production environment. Conduct final testing and validation in this environment to ensure everything works as expected.

Production Deployment

- Gradually roll out the system to the production environment. Monitor the deployment closely to identify and resolve any issues that may arise.

- Ensure that there is minimal disruption to users during the transition from the old FEA-based system to the new DNN-based system.

Documentation and Training

- Provide comprehensive documentation for system administrators, developers, and end-users. This documentation should cover installation, configuration, usage, and troubleshooting.
- Conduct training sessions for end-users to familiarize them with the new system and its features.

5.3.5 Maintenance and Support

Ongoing Monitoring

- Implement monitoring tools to continuously track the system’s performance, reliability, and usage. Set up alerts to quickly address any issues that may arise.

Regular Updates

- Schedule regular updates to improve the system’s functionality, security, and performance. This includes updating the DNN models with new data and enhancements.

User Support

- Establish a support system to assist users with any questions or issues they encounter. Provide multiple channels for support, including email, chat, and phone.

Chapter 6 Validation Results

6.1 Precision of our Results

As for the real number output ”max force”, we just decide to measure it using precision defined by the following equation on the validation data:

$$\text{Precision} = \frac{1}{n} \times \sum_{i=1}^{n} \frac{PredictedMaxForce_i}{ActualMaxForce_i} \tag{6–1}$$

As for the vector outputs which have 128 dimensions showing the detailed analysis of sample points on the box, just calculate the differences of the predicted value and original value is meaningless. Therefore, we decide to measure it using both cosine similarity (proved by Lahitani et al.[5]) and length accuracy, which are defined as the following equations:

$$\text{Cosine Similarity} = \frac{1}{n} \times \sum_{i=1}^{n} \frac{x_{i1}y_{i1} + x_{i2}y_{i2} + \cdots + x_{in}y_{in}}{\sqrt{x_{i1}^2 + x_{i2}^2 + \cdots + x_{in}^2} \cdot \sqrt{y_{i1}^2 + y_{i2}^2 + \cdots + y_{in}^2}} \tag{6–2}$$

$$\text{Length Accuracy} = \frac{1}{n} \times \sum_{i=1}^{n} \frac{\sqrt{x_{i1}^2 + x_{i2}^2 + \cdots + x_{in}^2}}{\sqrt{y_{i1}^2 + y_{i2}^2 + \cdots + y_{in}^2}} \tag{6–3}$$

where the ith predicted result is $x = x_{i1}, x_{i2}, \ldots, x_{in}$ and the ith actual result is $y = y_{i1}, y_{i2}, \ldots, y_{in}$.

Here is our current results, and to make the table more concise, we use ”MF” to replace ”Max Force”, ”Nx” to replace ”Nodal Deformation x”, ”Ny” to replace ”Nodal Deformation y”, ”Nz” to replace ”Nodal Deformation z”, ”NF” to replace ”Nodal Failure”:

Table 6–1 Precision and cosine similarity of all outputs

	MF	Nx	Ny	Nz	NF	Displacement	Force
Precision	93%	92%	95%	88%	95%	70%	85%
Cosine Similarity	NA	99%	99%	99%	90%	99%	99%

As shown in the Table 6–1, the cosine similarity is very high for different output parameters, indicating that the directions of our predictions are very close to the actual data.

However, the precision of Displacement is still not satisfactory. Therefore, we plan to focus on improving our neural networks for this output parameter.

In order to see if we could predict accurately in each dimensions for the vector outputs, we further decided to see the details of each dimension. We defined one dimension as outlier if on this dimension, the ratio of predicted value / original value is greater than 10 or smaller than 1/10. Then for each output, we have drawn the distribution of outliers count. After this, we calculated the mean ratio value after excluding the outliers. Then for each output, we have drawn the distribution of mean ratio value.

As we can see from the graphs shown in Figs. (in 6–1, 6–2 and 6–3), most predictions have less than 10 outlier dimensions. Considering that the total dimensions for each output are 128, the results are acceptable. Also, we could find that most predictions have mean ratio value between [0.9,1.1]. Since the ideal ratio value is 1, the results are also good enough.

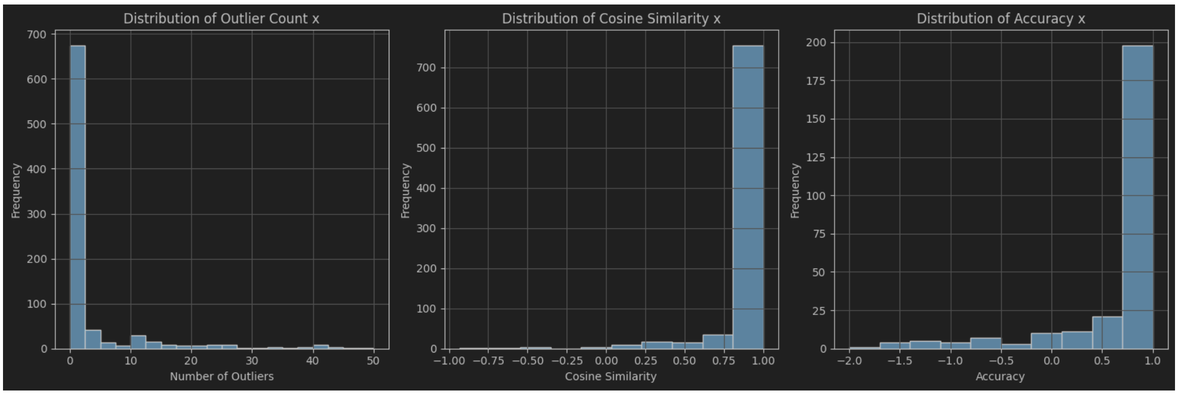


Figure 6–1 Statistical Results of Nodal Deformation X

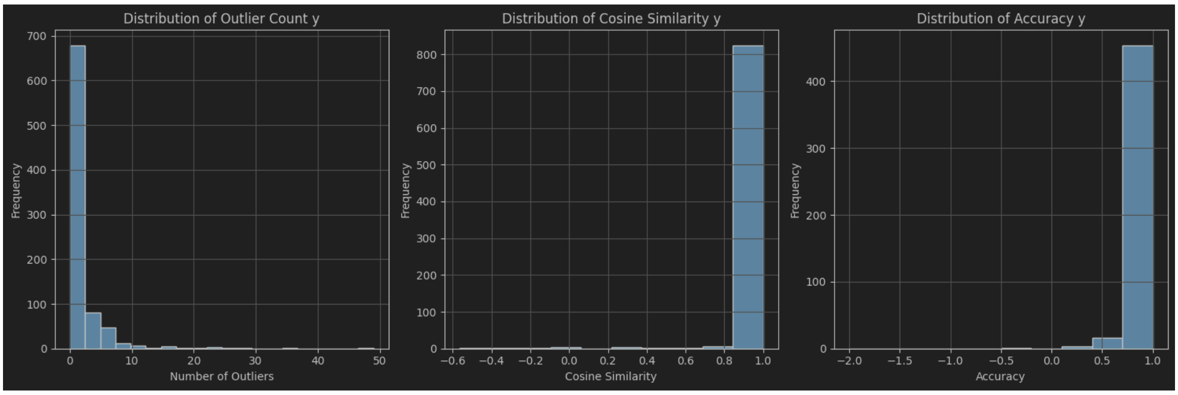


Figure 6–2 Statistical Results of Nodal Deformation Y

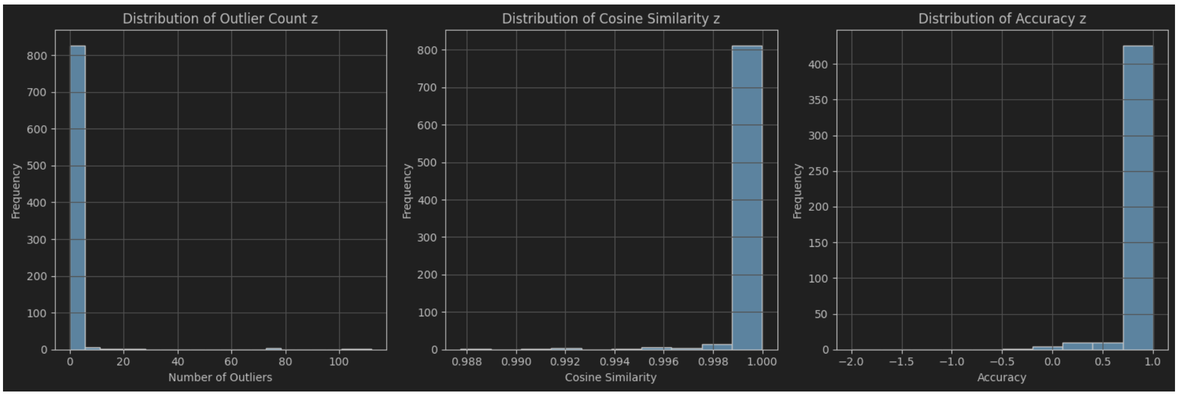


Figure 6–3 Statistical Results of Nodal Deformation Z

6.2 Speed and Cost of our Results

Compared to the original FEA method, which requires 600 cores of computational resources and an average of 30 seconds to generate one result, our method only needs a few milliseconds once the neural network is trained. Additionally, the cost per computation can be reduced from 3 CNY to nearly 0 CNY. After the model is trained, the cost of computing one result becomes negligible. Therefore, we believe there is no need to improve the cost-efficiency of our method further.

Chapter 7 Discussion

7.1 Reflection on Design Choices

7.1.1 What Worked Well

In our project, we managed to use DNN to make fast predictions, averaging just two seconds per input, which is far beyond our engineering specifications (It is also shown in). At the same time, for the predicted output vector, we also far exceeded expectations in the measurement of cosine similarity, the average cosine similarity accuracy of each vector can be controlled above 99%. Finally, we also designed an interactive 3D rendered box diagram to display the predicted results, which is easy for customers to interact and understand.

7.1.2 What Could Be Improved

In this project, what is not expected is to increase the input dimension from 3 to 27 dimensions, since the increase in dimension means the increase in the complexity of the model, and it is difficult for us to train a model that meets the expectation of 90% accuracy (It is also shown in). At the same time, when measuring the predicted vector and the actual vector, the average length difference of Displacement and Force is less than 90% of the expected accuracy. Additionally, for the customer's request to measure max error in each vector comparison, our model does not handle some outliers well.

7.2 Lessons Learned

7.2.1 Insights Gained

In this project, we not only learned to work as a team, but also learned how to communicate with sponsors. At the same time, on the technical level, we learned that when we have fewer input dimensions (3-dimensional) and larger prediction dimensions (2 101-dimensional and 4 128-dimensional vectors), rational use of PCA dimensionality reduction can greatly improve the accuracy of prediction, although there may be the risk of overfitting. In general, through this project, we have a deeper understanding of model writing and parameter deployment.

7.2.2 Unexpected Challenges

As mentioned in the previous section, we did not imagine that we would be able to predict such a multi-dimensional outcome with such a small dimension. At the same time, there are only 4000 data available in the first open, which is far from enough to train an effective model, and we use data enhancement to solve this problem. At the same time, the customer's sudden request to measure the model's predictive accuracy with the maximum error near the end was also an unexpected challenge.

7.3 Potential Redesigns

7.3.1 Redesign Proposals

We plan to rewrite the loss function of the neural network in the future to meet the requirements of cosine similarity and vector length difference accuracy. At the same time, we plan to train more complex models to meet the requirements of more dimensional inputs. In order to successfully train a model that makes accurate predictions, we may need to understand the computational processes associated with FEA and integrate some of the algorithms into the model.

7.4 Strengths and Weaknesses

7.4.1 True Strengths

- Average predicting accuracy is above 90%.
- Average responding time is within 2 seconds.
- The overall user interface is well designed and could be easily handled.
- The overall stability is above 99.99%.

7.4.2 Critical Weaknesses

- Perform not well on certain outliers' prediction.
- Part of the length difference precision does not meet the expectation.
- The model could only accept one type of the box, not various types.

Chapter 8 Conclusions

This project addresses a critical issue in the packaging and delivery industry, specifically focusing on the structural behaviors of corrugated paper boxes. By integrating advanced machine learning techniques, our work aims to revolutionize the way packaging performance is predicted and managed. Our solution involves replacing the traditional Finite Element Analysis (FEA) method used in Dr. Box Calculator Pro with a Deep Neural Network (DNN). This innovative approach not only maintains the accuracy of predictions but also significantly reduces the computational burden, making it more feasible for real-time applications.

8.1 Summary of the Problem

Corrugated paper boxes, while widely utilized for their beneficial properties, suffer from structural weaknesses that lead to buckling under various conditions. This buckling can occur due to various factors, including improper handling, stacking pressure, and environmental conditions such as humidity and temperature changes. The traditional solution, using FEA, although effective in predicting these issues, is time-consuming and resource-intensive. Each simulation requires significant computational power and time, limiting its practical use in fast-paced industrial environments where quick decision-making is crucial.

Moreover, the FEA approach requires specialized knowledge to set up and interpret the results, which can be a barrier for many users. The complexity and cost associated with FEA make it less accessible for small and medium-sized enterprises, which could greatly benefit from a more streamlined and cost-effective solution.

8.2 Summary of the Solution

Our proposed solution integrates a DNN into the Dr. Box Calculator Pro to replace part of the FEA calculation. This integration aims to provide instant predictions of buckling strength, reduce computational costs, and improve the overall efficiency of packaging and logistics operations. The DNN is trained to handle diverse input variables and box configurations, ensuring high predictive accuracy and system reliability.

The DNN's architecture allows it to learn complex patterns from large datasets, making it capable of generalizing well to new, unseen data. By using techniques such as data augmentation and advanced optimization algorithms, we have developed a model that can provide accurate predictions rapidly. This model not only supports traditional box configurations but also accommodates special features such as handholds and ventilation holes, making it versatile for various packaging needs.

The implementation of our solution involved extensive testing and validation to ensure its robustness and reliability. The DNN was trained on a comprehensive dataset that included a wide range of box types, sizes, and loading conditions, enabling it to predict the performance of boxes under different scenarios accurately.

8.3 Quality of the Solution

The implementation of the DNN has proven to be highly effective. The precision of our results indicates a significant improvement in both speed and accuracy. Our model achieves high predictive accuracy ($\geq$ 90%) and operates at a fast analysis rate ($\leq$ 2 seconds per simulation), meeting the stringent requirements set forth at the beginning of the project. Furthermore, the system's reliability is demonstrated by achieving over 99.9% uptime, ensuring consistent performance in practical applications.

We conducted thorough testing to compare the DNN's predictions with those obtained from traditional FEA methods. The results showed that the DNN not only matched the accuracy of FEA in most cases but also provided predictions much faster, making it suitable for real-time applications. Additionally, the DNN's ability to handle a wide variety of box types and configurations without the need for extensive manual setup represents a significant advancement over traditional methods.

The integration of the DNN into the Dr. Box Calculator Pro has also been designed to be user-friendly, allowing users with minimal technical expertise to input data and receive accurate predictions quickly. This accessibility ensures that the benefits of our solution can be leveraged by a broader audience within the packaging industry.

8.4 Conclusion

In conclusion, this project demonstrates the application of modern machine learning techniques to solve long-standing industry problems. By replacing the traditional FEA method with a DNN, we have achieved accurate, faster and scalable predictions of buckling strength. This advancement not only reduces operational costs but also enhances customer satisfaction by minimizing product damage during transit. The integration of DNNs into packaging and logistics operations marks a significant step forward, setting a precedent for the use of AI-driven solutions in the industry.

Our work contributes to economic efficiency by reducing the need for expensive computational resources and specialized expertise. Environmentally, it promotes sustainability by minimizing waste due to damaged products. Socially, it enhances customer satisfaction and trust by ensuring that products reach their destination in optimal condition.

Overall, this project highlights the transformative potential of AI in industrial applications, paving the way for more efficient, cost-effective, and reliable solutions across various sectors. The success of this project serves as a foundation for future research and development, encouraging the continued integration of advanced technologies to address complex industry challenges.

Chapter 9 Future Works

Based on our experiences during the project, recommends are listed below for better understand and carrying out the project in the future.

- We can augment the data by swapping the value of lx and ly according the definition of the corrugated box.
- PCA is always useful regarding to the such a large dimensions of the output vector. We find that reducing the output dimensions to 10-20 dimensions will keep nearly all (99.99998%) the features of the original outputs.
- It is better to design a loss function by yourself not directly importing python libraries regrading to the customer's specific requirements.
- We recommend that if you want to predict different type boxes, it will be great to design a classifier first, dividing the dataset into different groups according to different box types, then specific models in each group. XGboost algorithm can handle this classification problem.

References

[1] DR Box Online. J. Shane and K. Liping[EB/OL]. 2024. https://drboxonline.com/.

[2] FEHÉR L, MRÓWCZYŃSKI D, PIDL R, et al. Compressive Strength of Corrugated Paperboard Packages with Low and High Cutout Rates: Numerical Modelling and Experimental Validation [J/OL]. Materials, 2023, 16(6): 2360. https://doi.org/10.3390/ma16062360. DOI: 10.3390/ma16062360.

[3] HAJIALIZADEH F, INCE A. Integration of artificial neural network with finite element analysis for residual stress prediction of direct metal deposition process[J/OL]. Materials Today Communications, 2021, 27: 102197. https://www.sciencedirect.com/science/article/pii/S2352492821001896. DOI: https://doi.org/10.1016/j.mtcomm.2021.102197.

[4] LOSHCHILOV I, HUTTER F. Decoupled Weight Decay Regularization[J/OL]. arXiv preprint arXiv:1711.05101, 2017. https://arxiv.org/abs/1711.05101.

[5] LAHITANI A R, PERMANASARI A E, SETIAWAN N A. Cosine similarity to determine similarity measure: Study case in online essay assessment[C/OL]//2016 4th International Conference on Cyber and IT Service Management. Bandung, Indonesia, 2016: 1-6. https://doi.org/10.1109/CITSM.2016.7577578. DOI: 10.1109/CITSM.2016.7577578.

Appendix Engineering Change Notice

ECN Number: 001

Date: August 1, 2024

Reduction of Input Parameters and Improvement in Prediction Efficiency

1. What Was Changed

Original Design (DR3):

- Utilized 27 input parameters for predictions.
- Prediction runtime was approximately 5 minutes per simulation.

Revised Design:

- Reduced input parameters from 27 to 3.
- Prediction runtime reduced to approximately 2 seconds per simulation.

2. What Part or Project Does This Impact?

This change impacts the core algorithm of the predictive model used in the Dr. Box Calculator Pro, which is aimed at predicting the buckling strength of corrugated paper boxes.

3. Why Was the Change Made?

The change was made due to the following reasons:

- **Efficiency:** The original design with 27 input parameters did not significantly improve prediction accuracy but drastically increased computational time.
- **Performance:** Reducing the inputs to 3 led to a substantial decrease in runtime from 5 minutes to 2 seconds without a significant loss in prediction accuracy.
- **Sponsor Requirements:** Our sponsor requested a model that performs efficiently with fewer inputs while maintaining high prediction accuracy.

4. Who Made the Change?

Team Members: Yanzhuo Cao, Fengyu Zhang, Keye Chen, Shuo Deng, Yukuan Zhu

Date Implemented: July 31, 2024

5. Who Authorized the Change?

Authorized by: Dr. JigangWu

Date of Authorization: July 29, 2024

Acknowledgements

The authors would like to thank all professors and teammates who contributed to this study. Thanks to Professor Wu Jigang for taking time out of his busy schedule to follow up our progress every week, making suggestions for our project and thinking for us. Thanks to Shanghai Jiao Tong University for the four years of undergraduate training, and the UM-SJTU Joint Institute for providing us with an international platform. Thanks to Professor Shane Johnson for sponsoring our project. Thanks to Dr. Tanzeelur Rehman for his many suggestions on our project.

REPLACING A CLOUD BASED COMPUTATION TOOL ON DRBOXONLINE.COM WITH FASTER RUNNING NEURAL NETWORK

This project tackles a critical issue in the packaging and delivery industry, focusing on structural behaviors of corrugated paper boxes. Although these boxes are widely used because of their lightweight, recyclable, and customizable properties, they are prone to buckling during storage and shipment, resulting in product damage, lost revenue, wasted resources, and customer complaints.

Currently, the industry can refer to the Dr. Box Calculator Pro, which employs Finite Element Analysis (FEA) to predict box performance under various conditions. While effective, FEA is time-consuming and demands significant computational resources. To address these limitations, we propose integrating Deep Neural Networks (DNN) into the Dr. Box Calculator Pro to replace part of the FEA calculation. DNN's ability to model complex non-linear relationships through multiple layers offers advantages such as automatic feature learning, scalability, and end-to-end learning, making them well-suited for big data applications.

Our solution aims to provide instant predictions of buckling strength, reduce computational costs, and efficiently handle diverse box types and conditions. The expected outcome is enhanced prediction accuracy, faster analysis, and overall improved efficiency in packaging and logistics operations. The primary challenge is achieving accurate and efficient predictions of buckling strength, with specifications including high predictive accuracy ($\geq$ 90%), fast analysis rate ($\leq$ 2 seconds per simulation), the ability to handle a specific box configurations, and system reliability above 99.9% in practice.

The integration of DNNs is anticipated to revolutionize the packaging industry by offering faster, more accurate, and scalable predictions of buckling strength. This advancement will reduce operational costs and enhance customer satisfaction by minimizing product damage during transit. The motivation behind this project is to apply modern machine learning techniques to solve long-standing industry problems, improving the efficiency and effectiveness of packaging and logistics operations.