**One Improved Wasserstein GAN with Gradient

Penalty for Grain Consumption Prediction

^{Pei}Li, Chunhua Zhu^{*}

^{***PeiLi, Chunhua Zhu****
** *Abstract***—Prediction of grain consumption is crucial for** $\frac{1}{\text{network (CGAN) is introduced [6]. The CGAN incorporates}$ **
** *}* **banalty for Grain Consumption**
Penalty for Grain Consumption
Pei Li, Chunhua Zhu*
Pei Li, Chunhua Zhu*
analyzing the changing trend and balancing the grain supply
analyzing the changing trend and balancing the gr IAENG International Journal of Computer Science

One Improved Wasserstein GAN with Gradient

Penalty for Grain Consumption Prediction

Pei Li, Chunhua Zhu⁺

Perdiction model using conditional generative adversarial

Perd IAENG International Journal of Computer Science

Reflection

Penalty for Grain Consumption Prediction

Pei Li, Chunhua Zhu*

Pei Li, Chunhua Zhu*

Prediction of grain consumption is crucial for the prediction model using c Experiment of Computer Science

SSETStein GAN with Gradi

m Consumption Prediction

Pei Li, Chunhua Zhu*

prediction model using conditional generativ

prediction model using conditional generativ

prediction model using c

and ty for Grain Consumption
 and Consumption
 and demand in Consumption
 and demand in China. Recently, the use of generative
 and demand in China. Recently, the use of generative (GAN) is is
 and demand in Chi Penalty for Grain Consumption Pr

Pei Li, Chunhua Zhu^{*}

Pei Li, Chunhua Zhu^{*}

Pei Li, Chunhua Zhu^{*}

Pei Li, Chunhua Zhu^{*}

Perdiction model using consumption is crucial for

analyzing the changing trend and balanc **here are the model in the set of the control of the model attention of grain consumption is crucial for**
hastract—Prediction of grain consumption is crucial for metwork (CGAN) is in
and demand in China. Recently, the us **Example 19 and Series Collocated** Pei Li, Chunhua Zhu^{*}

Pei Li, Chunhua Zhu^{*}

Pei Li, Chunhua Zhu^{*}

prediction model us

analyzing the changing trend and balancing the grain supply

and demand in China. Recently, th **Pei Li, Chunhua Zhu***
 performance and the consumption is crucial for
 performance and stability, the use of generative
 performance and addemination in China. Recently, the use of generative
 adversarial networks Pei Li, Chunhua Zhu^{*}
 Pei Li, Chunhua Zhu^{}*
 prediction model using con
 Mbstract—Prediction of grain consumption is crucial for
 metwork (CGAN) is introduce

and demand in China. Recently, the use of generative Pet L₁, Chunhua Zhu^{*}

Mostract—Prediction of grain consumption is crucial for

prediction model using conce

analyzing the changing trend and balancing the grain supply

LSTM and multilayer percept

and demand in China *bstract***—Prediction of grain consumption is crucial for prediction model usin analyzing the changing trend and balancing the grain supply
and demand in China. Recently, the use of generative discriminator, respected
ave** *Hostract***—Prediction of grain consumption is crucial for network (CGAN) is in analyzing the changing trend ad balancing the grain supply and elemental in China. Recently, the use of generating future data has gained atte** *Abstract***—Prediction of grain consumption is crucial for** retwork (CGAN) is analyzing the changing trend and balancing the grain supply

and demand in China. Recently, the use of generative LSTM and multilaye

and demand *Abstract***—Prediction of grain consumption is crucial for reduction model assumption analyzing the changing trend and balancing the grain supply LSTM and multilayer perception and demand in China. Recently, the use of ge COLUT TEAT TERENT TERENT TERENT IS THE COLUTE THE CONTRACT TERM** and demand in China. Recently, the use of generative discriminator, respectived at a for generating future data has gained attention of information from tim **incorporates the mean square error (MSE)** between real and
 incorporates the mean square of generative discriminator, respectively
 intime-series prediction. In order to enhance prediction from daily the

in time-seri and usmand in Cuman. Fecentively, the use of generative discriminator, respectively
adversarial networks (GAN) to capture the distribution of information from daily to
in time-series prediction. In order to enhance predict **Loss function of the discriminator includes the L₁ norm as the discrimination of the L₂ norm used in the proposed WGAN was sense the discriming function. In order to enhance prediction chancing stock prediction can an gradient penalty term to enhance sparsity and resultation** (EMD-WGAN) for financial restaction and or generator and address model instability, an improved researchers have proposed v
Masserstein GAN with gradient penalt In thus-series precurson. In out to emante precurson
performance and address model instability, an improved researchers have propo
Wasserstein GAN with gradient penalty, referred to as financial time serious p
IWGAN-GP, is **EXPERTIFY ASSET THE THE CONDUCE IN THE CONDUCE IN THE CONDUCE TO HANGE THE CONDUCED AND AN INTERFERT OF AN AVERT CONDUCED AND AN INTERFERT OF AN AVERT CONDUCED IN A BORD AND A BUT CONSUMING THE CONSUMING THE CONSUMING THE 2020 Experimental results on gradient penalty, referred to assumption**
 EXAM-GP, is **proposed. The IWGAN-GP utilizes** a a Wasserstein GAN
 BILSTM) as the generator and a convolutional neural network with a GRU
 (CN Predictional long short-term memory neural tetwork and muck and the enterpredictional long short-term memory neural network with a GRU as the (CNN) as the discriminator, combining the memory capabilities of the server (C Buttecholar long short-term memory he (BiLSTM) as the generator and a convolutional i
(CNN) as the discriminator, combining the memor
of LSTM with the nonlinear feature extraction
CNN. Specifically, the loss function of
in *Index Terms***—grain consumption prediction, BiLSTM, CONSTAT WHAT HE HOMINICAT EXINT WHAT FORM CONSTANTS CONSTANT CONSTANT CONSTANT CONSTANT INCORPORATES the mean square error (MSE) between real argenerated samples to optimize the LSTM network, while the loss function of th** IT Square error (MSE) between real and

the optimize the LSTM network, while the

iscriminator includes the L₁ norm as the

it to enhance sparsity and robustness, in

rm used in existing WGAN-GP models.

(El

on grain c

EXEMIEV EXELE SUBATELY AND EXELE ANGLE AND EXELE ANGLE SUPPLANE CONSUMPTION
 EXECUTE ANGLE A EXECUTE AND
 EXECUTE THE CONDUCTION
 EXECUTE AND
 EXECUTE THE CONDUCTION
 CON-GP models.
 CON-GP mode **EXECT THE CONSUMISET AND**
 EXECT THE CONSUMISET AND *Index Terms*—grain consumption prediction, BiLSTM, $\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{2}$ (WGAN-GP, L₁ norm; EEMD

WGAN-GP, L₁ norm; EEMD

I. INTRODUCTION (WGAN-GP to address the
 $\frac{1}{2}$ CCURATELY understandin *Index Terms*—grain consumption prediction, BiLSTM, Furthermore, the EMD-
WGAN-GP, L₁ norm; EEMD
WGAN-GP to address
[11]. Shuntaro Takahas
I. INTRODUCTION time series model FIN
grain consumption is crucial for strengthe **WGAN-GP, L₁ norm; EEMD**

WGAN-GP to address the in
 \blacksquare I. INTRODUCTION

L. INT For the interest of a generator and a
discriminator characteristic of a generator of a generator of a generator of a generator of grain consumption is crucial for strengthening the macro
strategic control of grain and ensu I. INTRODUCTION

L. INTRODUCTION

L. INTRODUCTION

L. INTRODUCTION

discriminator combine MLP

discriminator combine MLP

shows potential advantages

strategic control of grain and ensuring the balance of grain

strategic **EXECURATELY understanding the evolving trend of** discriminator combine MLP and extragric control of grain and ensuring the balance of grain shows potential advantages is strategic control of grain and ensuring the balanc CURATELY understanding the evolving trend of shows potential advantages is trategic control of grain and ensuring the balance of grain is transparently and demand. Grain consumption data typically the demand of control of **The axample is at the series in the series and a** show including the matrix strategic control of grain and ensuring the balance of grain is also related to the character
follows a time series pattern. Among deep learning strategic control of grain and ensuring the balance of grain

supply and demand. Grain consumption data typically

follows a time series pattern. Among deep learning

follows a time series pattern. Among deep learning

tec mise in time series prediction. The emerging generative
versarial networks (GAN) performs better to explore time
ites prediction [6]. GAN consists of a generator and a
network (Bil
network (Bil
network (Bil
network (Bil
ne versarial networks (GAN) performs better to explore time

ires prediction [6]. GAN consists of a generator and a

incorporates bidirect

scriminator, which will be trained against each other [7].

AN has been effective in series prediction [6]. GAN consists of a generator and a
discriminator, which will be trained against each other [7].
GAN has been effective in image generation but nor yet in
time series prediction. Given the linearity an Series prediction [6]. GAN consists of a generator and a

discriminator, which will be trained against each other [7].

GAN has been effective in image generation but nor yet in

time series prediction. Given the linearity AN has been effective in image generation but nor yet in

ine series prediction. Given the linearity and stationarity of

square error (MSE)

inin consumption data, GAN for time series prediction is

coretically feasible. ime series prediction. Given the linearity and stationarity of grain consumption data, GAN for time series prediction if theoretically feasible. In 2021, an emotion-guided stock price Mericleanly feasible. In 2021, an emot

2022ZKCJ13.
Pei Li is an Associate Professor in Zhengzhou Tourism College, Experience Series prediction. Origin the initial of Satisfan and Satisfan consumption data, GAN for time series prediction is correctically feasible. In 2021, an emotion-guided stock price the discriminator loss correctic grain consumption data, GAN for time series prediction is

theoretically feasible. In 2021, an emotion-guided stock price

theoretically feasible. In 2021, an emotion-guided stock price
 $\frac{1}{1}$ norm instead of

Manuscr

zhuchunhua@haut.edu.cn).

EV and results are the manned of the discriminator includes the L₁ norm as the decomposition generative

to the L₂ norm used in existing WGAN-GP models.

The disc are proposed IWGAN-GP improves and multiple LSTM networ ent penalty term to enhance sparsity and robustness, in

and in existing WGAN-GP models.

Intensity and intersity and the Lamburg CEMD-WGAN) for financial

reinnental results on [gr](#page-0-0)ain consumption data from 1981 to

generat contrast to the L₂ norm used in existing WGAN-GP models.

Experimental results on grain consumption data from 1981 to
 Experimental results on grain consumption the proposed IWGAN-GP improves

and multiple LSTM networ $\mathop{\mathrm{ein}}\nolimits$ GAN with Gradient
 $\mathop{\mathrm{sum}}\nolimits_{\text{hua Zhu}^*}$
 $\mathop{\mathrm{prediction\ model\ using\ conditional\ generative\ adversarial\ network\ (CGAN)\ is\ introduced\ [6].}$ The CGAN incorporates

LSTM and multilayer perceptron (MLP) in the generator and

discriminator, respectively, along with $\mathop{\mathbf{e}}\limits$ eximple to $\mathop{\mathbf{GAN}}\limits_{\text{the number of the model that }T} \mathop{\mathbf{CCAN}}\limits_{\text{intra}}$ and multilayer perceptron (MLP) in the generator and discriminator, respectively, along with the emotional information from daily tweets as a condit **Example 18 CEN SUMPLE SUMPLE SUMPLE SUMPLE SUMPLE SUMPLE SUMPLE THE SUMPLE PROPRED MULTI PROPRED AND MULTI AND Example 18 CONSTAN With Gradient**
 Example 18 Constant Summary Summary Constant Summary along with the emotional information, respectively, along w SUMPTION Prediction

mediation from data as a conditional generative adversarial

prediction model using conditional generative adversarial

network (CGAN) is introduced [6]. The CGAN incorporates

LSTM and multilayer p **Example 18 Constrained Strate Procediction**

thua Zhu^{*}

prediction model using conditional generative adversarial

network (CGAN) is introduced [6]. The CGAN incorporates

LSTM and multilayer perceptron (MLP) in the gen **EXAMPLE COLOMPT PECATCLE COLOM**

hua Zhu^{*}

prediction model using conditional generative adversarial

network (CGAN) is introduced [6]. The GGAN incorporates

LSTM and multilayer perceptron (MLP) in the generator and

d Financial time serious conditional generative adversarial
financism and multilayer perceptron (MLP) in the generator and
discriminator, respectively, along with the emotional
find discriminator, respectively, along with th hua Zhu^{*}
prediction model using conditional generative adversarial
network (CGAN) is introduced [6]. The CGAN incorporates
LSTM and multilayer perceptron (MLP) in the generator and
discriminator, respectively, along with hua Zhu^{*}
prediction model using conditional generative adversarial
network (CGAN) is introduced [6]. The CGAN incorporates
LSTM and multilayer perceptron (MLP) in the generator and
discriminator, respectively, along with prediction model using conditional generative adversarial
network (CGAN) is introduced [6]. The CGAN incorporates
LSTM and multilayer perceptron (MLP) in the generator and
discriminator, respectively, along with the emotio prediction model using conditional generative adversarial
network (CGAN) is introduced [6]. The CGAN incorporates
LSTM and multilayer perceptron (MLP) in the generator and
discriminator, respectively, along with the emotio prediction model using conditional generative adversarial
network (CGAN) is introduced [6]. The CGAN incorporates
LSTM and multilayer perceptron (MLP) in the generator and
discriminator, respectively, along with the emotio network (CGAN) is introduced [6]. The CGAN incorporates
LSTM and multilayer perceptron (MLP) in the generator and
discriminator, respectively, along with the emotional
information from daily tweets as a conditional input,
 LSTM and multilayer perceptron (MLP) in the generator and
discriminator, respectively, along with the emotional
information from daily tweets as a conditional input,
enhancing stock prediction accuracy. Additionally,
resea discriminator, respectively, along with the emotional
information from daily tweets as a conditional input,
enhancing stock prediction accuracy. Additionally,
researchers have proposed various GAN-based models for
financia information from daily tweets as a conditional input,
enhancing stock prediction accuracy. Additionally,
researchers have proposed various GAN-based models for
financial time serious prediction. Lin H. C et al. [8] propose enhancing stock prediction accuracy. Additionally,
researchers have proposed various GAN-based models for
financial time serious prediction. Lin H. C et al. [8] proposes
a Wasserstein GAN with gradient penalty (WGAN-GP)
ne researchers have proposed various GAN-based models for
financial time serious prediction. Lin H. C et al. [8] proposes
a Wasserstein GAN with gradient penalty (WGAN-GP)
network with a GRU as the generator and a convolution financial time serious prediction. Lin H. C et al. [8] proposes
a Wasserstein GAN with gradient penalty (WGAN-GP)
network with a GRU as the generator and a convolutional
neural network (CNN) as the discriminator to investi a Wasserstein GAN with gradient penalty (WGAN-GP)
network with a GRU as the generator and a convolutional
neural network (CNN) as the discriminator to investigate
whether adversarial systems can help improve time series
pr network with a GRU as the generator and a convolutional
neural network (CNN) as the discriminator to investigate
whether adversarial systems can help improve time series
prediction performance. Experiments have demonstrate neural network (CNN) as the discriminator to investigate
whether adversarial systems can help improve time series
prediction performance. Experiments have demonstrated that
the adversarial network outperforms the tradition whether adversarial systems can help improve time series
prediction performance. Experiments have demonstrated that
the adversarial network outperforms the traditional LSTM
network. Wang Jing et al. [9] introduces an empir prediction performance. Experiments have demonstrated that
the adversarial network outperforms the traditional LSTM
network. Wang Jing et al. [9] introduces an empirical mode
decomposition generative adversarial network
(E the adversarial network outperforms the traditional LSTM
network. Wang Jing et al. [9] introduces an empirical mode
decomposition generative adversarial network
(EMD-WGAN) for financial time series prediction. The
generato network. Wang Jing et al. [9] introduces an empirical mode
decomposition generative adversarial network
(EMD-WGAN) for financial time series prediction. The
generator consists of empirical mode decomposition (EMD)
and mult decomposition generative adversarial network
(EMD-WGAN) for financial time series prediction. The
generator consists of empirical mode decomposition (EMD)
and multiple LSTM networks, while the discriminator adopts
CNN. The (EMD-WGAN) for financial time series prediction. The generator consists of empirical mode decomposition (EMD) and multiple LSTM networks, while the discriminator adopts CNN. The data after EMD exhibit similar frequency and generator consists of empirical mode decomposition (EMD)
and multiple LSTM networks, while the discriminator adopts
CNN. The data after EMD exhibit similar frequency and
good regularity, reducing the complexity of the gene and multiple LSTM networks, while the discriminator adopts
CNN. The data after EMD exhibit similar frequency and
good regularity, reducing the complexity of the generated
model and enhancing the prediction accuracy [10].
F CNN. The data after EMD exhibit similar frequency and
good regularity, reducing the complexity of the generated
model and enhancing the prediction accuracy [10].
Furthermore, the EMD-WGAN utilizes the loss function of
WGAN od regularity, reducing the complexity of the generated
odel and enhancing the prediction accuracy [10].
rthermore, the EMD-WGAN utilizes the loss function of
GAN-GP to address the instability of the original GAN
1]. Shunt model and enhancing the prediction accuracy [10].
Furthermore, the EMD-WGAN utilizes the loss function of
WGAN-GP to address the instability of the original GAN
[11]. Shuntaro Takahashi et al. [12] proposes the financial
t Furthermore, the EMD-WGAN utilizes the loss function of WGAN-GP to address the instability of the original GAN [11]. Shuntaro Takahashi et al. [12] proposes the financial time series model FIN-GAN, where the generator and WGAN-GP to address the instability of the original GAN
[11]. Shuntaro Takahashi et al. [12] proposes the financial
time series model FIN-GAN, where the generator and
discriminator combine MLP and CNN. Therefore, GAN
shows CCURATELY understanding the evolving trend of
shows potential advantages in time series prediction.
strategic control of grain and ensuring the balance of grain
less function at a base of marine the series prediction,
less

[11]. Shuntaro Takahashi et al. [12] proposes the financial
time series model FIN-GAN, where the generator and
discriminator combine MLP and CNN. Therefore, GAN
shows potential advantages in time series prediction.
However time series model FIN-GAN, where the generator and
discriminator combine MLP and CNN. Therefore, GAN
shows potential advantages in time series prediction.
However, for specific applications, the network structure,
loss fun discriminator combine MLP and CNN. Therefore, GAN
shows potential advantages in time series prediction.
However, for specific applications, the network structure,
loss function, etc., have an impact on prediction performan shows potential advantages in time series prediction.
However, for specific applications, the network structure,
loss function, etc., have an impact on prediction performance,
which is also related to the characteristics o However, for specific applications, the network structure,
loss function, etc., have an impact on prediction performance,
which is also related to the characteristics of nonlinearity and
non-stationarity in historical data loss function, etc., have an impact on prediction performance,
which is also related to the characteristics of nonlinearity and
non-stationarity in historical data series.
To enhance prediction performance and address mode which is also related to the characteristics of nonlinearity and
non-stationarity in historical data series.
To enhance prediction performance and address model
instability in grain consumption prediction, an improved
WGAN rediction performance and address model
grain consumption prediction, an improved
IWGAN-GP) is proposed. The IWGAN-GP
pidirectional long short-term memory neural
LSTM) as generators and CNN as the
with a new loss function stability in grain consumption prediction, an improved GAN-GP (IWGAN-GP) is proposed. The IWGAN-GP corporates bidirectional long short-term memory neural twork (BiLSTM) as generators and CNN as the scriminator, with a new WGAN-GP (IWGAN-GP) is proposed. The IWGAN-GP
incorporates bidirectional long short-term memory neural
network (BiLSTM) as generators and CNN as the
discriminator, with a new loss function introducing mean
square error (MS incorporates bidirectional long short-term memory neural
network (BiLSTM) as generators and CNN as the
discriminator, with a new loss function introducing mean
square error (MSE) to optimize the generator and ensure
gener network (BiLSTM) as generators and CNN as the discriminator, with a new loss function introducing mean square error (MSE) to optimize the generator and ensure generated data closely resemble real data. Furthermore, the di discriminator, while a licew loss function introd
square error (MSE) to optimize the generator
generated data closely resemble real data. Furtl
discriminator loss function in WGAN-GP is mo
L1 norm instead of L2 norm, enha a hew loss function infoducing incarries
to optimize the generator and ensure
ly resemble real data. Furthermore, the
notion in WGAN-GP is modified to use
of L2 norm, enhancing sparsity and
ucing model complexity.
IE PROPO

 $\tilde{\mathbf{x}}_{t+1}$ at time $t+1$; then, it contains

 \tilde{x}_{t+1} to produce fake an Fig. 1 The proposed IWGAN-GP architecture.

Fig. 1 The proposed IWGAN-GP architecture.

Fig. 1 The proposed IWGAN-GP architecture.

The series information between x_{t-2} , x_{t-1} , x_t , x_{t+1} , similarly, real data $\tilde{\chi}_{t+1}$ Fig. 1 The proposed IWGAN-GP architecture.

Fig. 1 The proposed IWGAN-GP architecture.
 x_{i-2}, x_{i-1}, x_i , \tilde{x}_{t+1} similarly, real data x_{i-2}, x_{i-1}, x_i , \tilde{x}_{t+1} is produce fake and maximum value data $x_{i-2}, x_{i-1},$ Fig. 1 The proposed IWGAN-GP architecture.
 x_{i-2}, x_{i-1}, x_i at the first three points with \overline{x}_{i+1} to produce fake and maximum va

data x_{i-2}, x_{i-1}, x_i , \overline{x}_{i+1} . similarly, real data $x_{i-2}, x_{i-1}, x_i, x_{i+1}$ is
 Fig. 1 The proposed IWGAN-GP architecture.
 x_{t-2}, x_{t-1}, x_t at the first three points with \tilde{x}_{t+1} to produce fake and maximum value of

data $x_{t-2}, x_{t-1}, x_t, \tilde{x}_{t+1}$. similarly, real data $x_{t-2}, x_{t-1}, x_t, x_{t+1}$ i Fig. 1 The proposed IWGAN-GP architecture.
 $x'_{t-2}, x'_{t-1}, x'_{t}$ at the first three points with \tilde{x}'_{t+1} to produce fake and maximum value of $X' = (x'_1, x'_2,..., x'_n)$ is the normal

data $x'_{t-2}, x'_{t-1}, x'_{t}$, similarly, x_{i-2}, x_{i-1}, x_i , x_{i+1} , similarly, real data $x_{i-2}, x_{i+1}, x_i, x_{i+1}$ is the 1
data $x_{i-2}, x_{i-1}, x_i, x_{i+1}$, similarly, real data $x_{i-2}, x_{i-1}, x_i, x_{i+1}$ is the 1
denoted as data at time *t*. Both fake data and real da data $x'_{t-2}, x'_{t-1}, x'_{t}, x'_{t+1}$. similarly, real data x'_{t-1}
denoted as data at time *t*. Both fake data are inputted into the discriminator to capture the
time series information between x'_{t+1}
 $x'_{t-2}, x'_{t-1}, x'_{$ denoted as data at time *t*. Both fake data and real data
inputted into the discriminator to capture the correlation
time series information between x_{i+1} , x_{i+1} , x_{i+2} , x_{i-1} , x_i . The discriminator of IWGAN outted into the discriminator to capture the correlation and

multiple empirical mode de-

are series information between x_{i+1} , x_{i+1} and noise to adaptively decompo-
 x_{i-1}, x_i . The discriminator of IWGAN-GP empl time series information between x_{i+1} , x_{i+1} and moise to adaptively decompose
 x_{i-2}, x_{i-1}, x_i . The discriminator of IWGAN-GP employs CNN, mixing. The EEMD decompositions where its output value Q represents the di Statistical Version of IWGAN-GP employs CNN, mixing the BEMD decomposition
 x_{r-2}, x_{r-1}, x_r . The discriminator of IWGAN-GP employs CNN, mixing sequence X' and

real data and generated data. The generator and discriminato

 x_{i-2}, x_{i-1}, x_i . The discriminator of IWGAN-GP employs CNN, mixing. The ELEWID decomposition where its output value Q represents the discrepancy between real data and generated data. The generator and discriminator times where its output value Q represents the discrepancy between
real data and generated data. The generator and discriminator
are trained through alternating iterations until the output
are trained through alternating iterati real data and generated data. The generator and discriminator

are trained through alternating iterations until the output

value of the discriminator converges close to zero or

fluctuates slightly around zero. The "+" i are trained through alternating iterations until the output

value of the discriminator converges close to zero or

to sequence X' the

value of the discriminator converges close to zero or

to sequence X' the

value of t value of the discriminator converges close to zero or

fluctuates slightly around zero. The "+" in Figure 1 signifies $X'_j = X' + 1$

vector concatenation.

A. Stabilization processing

The dataset used in this study is sour fluctuates slightly around zero. The "+" in Figure 1 signifies $X_j = 2$
vector concatenation.
A. *Stabilization processing* 3) Decomposing X
vector concatenation.
A. *Stabilization processing* 3) Decomposing $X_j = 2$
vector vector concatenation.

A. Stabilization processing

The dataset used in this study is sourced from the official

website of the National Bureau of Statistics and the China

Statistical Yearbook. China's grain consumption 21. Stabilization and the statest unit in the decomposing $X_1, X_2, ..., A_n$.

The dataset used in this study is sourced from the official

Statistical Yearbook. China's grain consumption is

Statistical Yearbook. China's grai A. Stabilization processing

The dataset used in this study is sourced from the official

website of the National Bureau of Statistics and the China

Statistical Yearbook. China's grain consumption is

Statistical Yearboo The dataset used in this study is sourced from the official
website of the National Bureau of Statistics and the China
Statistical
Yearbook. China's grain consumption is
4) Averaging the correspon-
cood grain including ra website of the National Bureau of Statistics and the China

Statistical Yearbook. China's grain consumption is

Statistical Yearbook. China's grain consumption is

4) Averaging the corrected prices and the composition and Statistical Yearbook. China's grain consumption is 4) Averaging
categorized into food grain and non-food grain [13], with perform
food grain including rations and feed grains, and non-food
grain including industrial grain tegorized into food grain and non-food grain [13], with

order grain including rations and feed grains, and non-food

in including industrial grain, seed grain and loss grain.
 $a_i = \frac{1}{m} \sum_{j=1}^m a_{i,j} (i = 1,$

the trend food grain including rations and feed grains, and non-food

grain including industrial grain, seed grain and loss grain.

The trend chard repicting China's total grain consumption

from 1981 to 2020 is illustrated in Figu grain including industrial grain, seed grain and loss grain.

The trend chart depicting China's total grain consumption

the processed from 1981 to 2020 is illustrated in Figure 2, with the ordinate

unit being 10,000 ton The trend chart depicting China's total grain consumption
from 1981 to 2020 is illustrated in Figure 2, with the ordinate
decomposition. As shown in
unit being 10,000 tons. The stationarity of grain consumption
derained b

from 1981 to 2020 is illustrated in Figure 2, with the ordinate
unit being 10,000 tons. The stationarity of grain consumption
unit root est method, with the calculated significance test
statistics indicating non-stationar unit being 10,000 tons. The stationarity of grain consumption

data is evaluated using the Augmented Dickey-Fuller (ADF)

unit root test method, with a calculated significance test

statistics indicating non-stationary wi data is evaluated using the Augmented Dickey-Fuller (ADF)

unit root test method, with the calculated significance test

tatistics indicating non-stationary with a P value of 0.887
 $\mathbf{E} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4$ unit root test method, with the calculated significance test
statistics indicating non-stationary with a P value of 0.887 $\mathbf{E} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5)^T_{n^5} = (\mathbf{e}_1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_3, \mathbf{$ statistics indicating non-stationary with a P value of 0.887 $E = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5)_{n^5} = (\frac{14}{3})$. Globally, the grain consumption displays the considered as the five feat characteristics of non-line [14]. Globally, the grain consumption displays the considered as the five features a characteristics of non-linearity, non-stationary, and an overall inputting into the generator, thes increasing trend, as evident from Fi characteristics of non-linearity, non-stationary, and an overall
increasing itend, as evident from Figure 2.
into three dimensions, inclusion contents are non-linearity and non-stationary of the
grain consumption data, st pletiaring Unitar solvat para consumption

is illustrated in Figure 2, with the ordinate

ording the Augmented Dickey-Fuller (ADF)

in the calculated significance test

ording the Augmented Dickey-Fuller (ADF)

undergoes The Augmented Dickey-Fuller (ADF) undergoos EEMD to obtain
the stationarity of grain consumption
the Augmented Dickey-Fuller (ADF) undergoos EEMD to obtain
the calculated significance test
that the calculated significance data is evaluated using the Augmented Dickey-Fuller (ADF)

unstigated component RE

unit root test methods, with the calculated significance test methods, with a Park of the data

tratistics indicating non-stationary with stationary time series refers to patterns in the series that

remain constant over time, which are essential for subsequent

prediction. Normalization and ensemble empirical mode

decomposition (EEMD) are employed to smoo

$$
X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}
$$

 \tilde{x}_{t+1} and noise to adaptively only **1) 1 D** the EEMD algorithm, proposed by Huang [15], involves
iple empirical mode decomposition with Gaussian white
to adaptively decompose the signal and avoid the mode
ng. The EEMD decomposition process includes:
Providing seque Maximum value of X, respectively, and

maximum value of X, respectively, and
 $(x_i, x_2,...,x_n)$ is the normalized series. Considering the

stationarity in X', EEMD is used to smooth X'.

Le EEMD algorithm, proposed by Huang [*m* value of *X*, respectively, and
 x_n) is the normalized series. Considering the
 x_n) is the normalized series. Considering the

algorithm, proposed by Huang [15], involves

ical mode decomposition with Gaussian

- times *m*;
-

$$
X'_{j} = X' + W_{j} \quad (j = 1, 2, ..., m)
$$
 (2)

- ise to adaptively decompose the signal and avoid the mode

xing. The EEMD decomposition process includes:

1) Providing sequence X' and the number of processing

times m;

2) Adding m groups of random white noise W_1, W_2 The EEMD decomposition process includes:

widing sequence X' and the number of processing

nes m;

ding m groups of random white noise $W_1, W_2, ..., W_m$

sequence X' to form $X_1, X_2', ..., X_m'$;
 $X_j' = X' + W_j$ ($j = 1, 2, ..., m$) (2)

co *i i* groups of random white noise $W_1, W_2, ..., W_m$

ince X' to form $X'_1, X'_2, ..., X'_m$;
 $X'_j = X' + W_j$ $(j = 1, 2, ..., m)$ (2)

osing $X'_1, X'_2, ..., X'_m$ using EMD to obtain a
 i Intrinsic Mode Function (IMF) components
 $\ldots, a_{i,m}$;

-

$$
a_i = \frac{1}{m} \sum_{j=1}^{m} a_{i,j} (i = 1, 2, \dots, N; j = 1, 2, \dots, m)
$$
 (3)

and the ormalized series. Considering the
 X' , EEMD is used to smooth X' .

Hen, proposed by Huang [15], involves

node decomposition with Gaussian white

elecompose the signal and avoid the mode

decomposition process *a a*, m , is the normalized series. Considering the
 a m , x_n , is the normalized series. Considering the

rity in X' , EEMD is used to smooth X' .

D algorithm, proposed by Huang [15], involves

pirical mode deco 2) Adding *m* groups of random white noise $W_1, W_2, ..., W_m$

to sequence *X* is to form $X'_1, X'_2, ..., X'_m$;
 $X'_j = X' + W_j$ ($j = 1, 2, ..., m$) (2)

3) Decomposing $X'_1, X'_2, ..., X'_m$ using EMD to obtain a

series of Intrinsic Mode Function to sequence X' to form $X_1, X_2, ..., X_m$;
 $X'_j = X' + W_j$ ($j = 1, 2, ..., m$) (2)

3) Decomposing $X'_1, X'_2, ..., X'_m$ using EMD to obtain a

series of Intrinsic Mode Function (IMF) components
 $a_{i,1}, a_{i,2}, ..., a_{i,m}$;

4) Averaging the cor $X_j = X' + W_j$ $(j = 1, 2, ..., m)$ (2)

3) Decomposing $X_1, X_2, ..., X_m'$ using EMD to obtain a

series of Intrinsic Mode Function (IMF) components
 $a_{i,1}, a_{i,2}, ..., a_{i,m}$;

4) Averaging the corresponding IMF components to

perform EEMD mixing. The EEMD decomposition process includes:

1) Providing sequence X' and the number of processing

times m;

2) Adding m groups of random white noise $W_1, W_2, ..., W_m$

to sequence X' to form $X_1, X_2, ..., X_m$;
 $X_j = X' + W_j$ (IMF) components

(IMF) components

to
 $i = 1, 2, ..., m$) (3)

components after

3, the sequence X'

components and one
 $N = 5$. Let

then e_n can be

ed with x'_n . Before

need to be reshaped

number of samples, $a_{i,1}, a_{i,2},..., a_{i,m}$;

4) Averaging the corresponding IMF components to

perform EEMD decomposition.
 $a_i = \frac{1}{m} \sum_{j=1}^{m} a_{i,j} (i = 1, 2,..., N; j = 1, 2,..., m)$ (3)

where *N* is the number of IMF components after

decomposition. 4) Averaging the corresponding IMF components to
perform EEMD decomposition.
 $a_i = \frac{1}{m} \sum_{j=1}^{m} a_{i,j} (i = 1, 2, ..., N; j = 1, 2, ..., m)$ (3)
where N is the number of IMF components after
decomposition. As shown in Figure 3, the se 4) Averaging the corresponding livit components to
perform EEMD decomposition.
 $a_i = \frac{1}{m} \sum_{j=1}^{m} a_{i,j} (i = 1, 2, ..., N; j = 1, 2, ..., m)$ (3)
where N is the number of IMF components after
decomposition. As shown in Figure 3, the perform EEWD decomposition.
 $a_i = \frac{1}{m} \sum_{j=1}^{m} a_{i,j} (i = 1, 2, ..., N; j = 1, 2, ..., m)$ (3)

where N is the number of IMF components after

decomposition. As shown in Figure 3, the sequence X'

undergoes EEMD to obtain four IMF c $a_i = \frac{1}{m} \sum_{j=1}^{m} a_{i,j} (i = 1, 2, ..., N; j = 1, 2, ..., m)$ (3)
where N is the number of IMF components after
decomposition. As shown in Figure 3, the sequence X'
undergoes EEMD to obtain four IMF components and one
residual compo where N is the number of IMF components after
decomposition. As shown in Figure 3, the sequence X'
undergoes EEMD to obtain four IMF components and one
residual component RES, so $N = 5$. Let
 $\mathbf{E} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_$ where *N* is the number of IMF components aft decomposition. As shown in Figure 3, the sequence *J* undergoes EEMD to obtain four IMF components and or residual component RES, so $N = 5$. In $\mathbf{E} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$

Volume 51, Issue 8, August 2024, Pages 1105-1111

Fig. 3 The stabilization processing

Fig. 3 The stabilization processing

B. The Timing generator based on BiLSTM and type of recurrent neural network (RNN) that is able to retain the reverse sequence. The

information fr Fig. 3 The stabilization processing

Fig. 3 The stabilization processing

B. The Timing generator based on BiLSTM

As previously mentioned, the generator utilizes BiLSTM,

a single-layer BiLSTM essence. The outright of th Fig. 3 The stabilization processing

Fig. 3 The stabilization processing

B. The Timing generator based on BiLSTM A single-layer BiLSTM As previously mentioned, the generator utilizes BiLSTM, the reverse sequence. The or
 Fig. 3 The stabilization processing

B. The Timing generator based on BiLSTM

As previously mentioned, the generator utilizes BiLSTM,

a type of recurrent neural network (RNN) that is able to retain

information from previ Fig. 3 The stabilization processing

B. The Timing generator based on BiLSTM

As previously mentioned, the generator utilizes BiLSTM,

a type of recurrent neural network (RNN) that is able to retain

information from previ Fig. 3 The stabilization processing
 B. The Timing generator based on BiLSTM

As previously mentioned, the generator utilizes BiLSTM,

ary one processes the forward sequence. The output

ary ary of recurrent neural netw *B. The Timing generator based on BiLSTM* A single-layer BiLSTM os

As previously mentioned, the generator utilizes BiLSTM,

and the reverse sequence. The out

information from previous time points to calculate the

infor *B. The Timing generator based on BiLSTM*
 EXECUTE: The processes the forward sequestion they as the generator utilizes BiLSTM, the reverse sequence. The output

any processes the forward seques and the area of recurren As previously mentioned, the generator utilizes BiLSTM,

strevers sequence. The outp

and type of recurrent neural network (RNN) that is able to retain

information from previous time points to calculate the

information by the of recurrent neural network (RNN) that is able

formation from previous time points to calcumation at the current time point [16]. RNNs have

at performance in handling time series data. I

NNs suffer from a signif **EXECUTE SET AN ANDEL CONTA C** Fig. 3 The stabilization processing
 B. The Timing generator based on BiLSTM

³¹ (see the carrelat text) (NNN) that is able to retain the concileration of the observation of the content term and province (RNN) that is ming generator based on BH.5TM

on processes the forward sequence, and the course

circular constructions. The reverse sequence. The culture of the two first

certificts content that is able to retain

constetement conste information at the current time point [16]. RNNs have shown
greetage ant
great performance in handling time series data. However,
These two c
RNNs suffer from a significant long-term dependency issue output of the
[3], w

 \mathbf{f}_t , the input fate \mathbf{i}_t and the out

$$
\mathbf{f}_{t} = \sigma(\mathbf{W}_{f}[\mathbf{h}_{t-1}; \mathbf{k}_{t}] + \mathbf{b}_{f})
$$
 (4)

$$
\mathbf{i}_{t} = \sigma(\mathbf{W}_{i}[\mathbf{h}_{t-1}; \mathbf{k}_{t}] + \mathbf{b}_{i}) \qquad (5)
$$

$$
\mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}; \mathbf{k}_t] + \mathbf{b}_o)
$$
 (6)

$$
\mathbf{s}_{t} = \mathbf{f}_{t} \mathbf{G}_{t-1} + \mathbf{i}_{t} \text{Ctanh}(\mathbf{W}_{s}[\mathbf{h}_{t-1}; \mathbf{k}_{t}] + \mathbf{b}_{s})
$$
 (7)

$$
\mathbf{h}_t = \mathbf{o}_t \ \text{Ctanh}(\mathbf{s}_t) \tag{8}
$$

state h_{t-1} and the current input k_t ; s_t is a memory cell at time t ; W_f , W_i , W_o , W_s and b_f , b_i , b_o , b_s are learning handling data $f_t = \sigma(W_f[h_{t-1}; k_t] + b_f)$ (4) widely
 $i_t = \sigma(W_f[h_{t-1}; k_t] + b_i)$ (5) convolu
 $o_t = \sigma(W_o[h_{t-1}; k_t] + b_o)$ (6) shapes
 $s_t = f_t \Im_{t-1} + i_t \text{Canh}(W_s[h_{t-1}; k_t] + b_s)$ (7) features
 $h_t = o_t \text{Canh}(s_t)$ (8) turn consum

where $[h_{t-1}; k_t]$ is a conca $\mathbf{o}_t = \sigma(\mathbf{W}_o[\mathbf{h}_{t-1}; \mathbf{k}_t] + \mathbf{b}_o)$ (6)
 $\mathbf{s}_t = \mathbf{f}_t \odot_{t-1} + \mathbf{i}_t \text{Ctanh}(\mathbf{W}_s[\mathbf{h}_{t-1}; \mathbf{k}_t] + \mathbf{b}_s)$ (7)
 $\mathbf{h}_t = \mathbf{o}_t \text{Ctanh}(\mathbf{s}_t)$ (8)

where $[\mathbf{h}_{t-1}; \mathbf{k}_t]$ is a concatenation of the previo

generates an output vector after three-time steps, while the **Example 19**
 Examples
 Ex Example 19 and 19 ^{and} S of features

and Bingle-layer BiLSTM essentially consists of two LSTMs:

and single-layer BiLSTM essentially consists of two LSTMs:

one processes the forward sequence, and the other processes

the reverse sequence The generator includes a BiLSTM layer with 128 neurons,

a single-layer BiLSTM essentially consists of two LSTMs:

e processes the forward sequence, and the other processes

reverse sequence. The outputs of the two LSTMs a Followed by two Dense layers with 64 and 1 neurons,
single samples
the reverse sequence. The outputs of the two LSTMs are then
concatenated. As illustrated in Figure 4, the forward LSTM
generates an output vector after th A single-layer BiLSTM essentially consists of two LSTMs:
one processes the forward sequence, and the other processes
the reverse sequence. The outputs of the two LSTMs are then
concatenated. As illustrated in Figure 4, th A single-layer BiLSTM essentially consists of two LSTMs:
one processes the forward sequence, and the other processes
the reverse sequence. The outputs of the two LSTMs are then
concatenated. As illustrated in Figure 4, th A single-layer BiLSTM essentially consists of two LSTMs:
one processes the forward sequence, and the other processes
the reverse sequence. The outputs of the two LSTMs are then
concatenated. As illustrated in Figure 4, th **Example 18 (Example)**
 Examples

A single-layer BiLSTM essentially consists of two LSTMs:

one processes the forward sequence, and the other processes

the reverse sequence. The outputs of the two LSTMs are then

conca

concatenated. As illustrated in Figure 4, the forward LST
generates an output vector after three-time steps, while t
reverse LSTM produces another output after three-time step
These two output vectors are combined to form nerates an output vector after three-time steps, while the
verse LSTM produces another output after three-time steps.
ese two output vectors are combined to form the final
tput of the BiLSTM [18, 19].
The generator includ reverse LSTM produces another output after three-time steps.
These two output vectors are combined to form the final
output of the BiLSTM [18, 19].
The generator includes a BiLSTM layer with 128 neurons,
followed by two D

at nework (KNN) that is able to retain

and the constrained and put vector after the point is colculate the

vious time point [16]. RNNs have shown reverse LSTM produces another out

and thigher the reserved and the dista Extramely through the concess a gas structure to regulate the

tring of features, including the forget gate
 \mathbf{i}_t and the output gate \mathbf{o}_t [17]. The LSTM C. CNN-based Discriminal

marized as follows:
 $\mathbf{f}_t = \sigma(\$ d forgetting of features, including the forget gate

ut fate \mathbf{i}_t and the output gate \mathbf{o}_t [17]. The LSTM C. CNN-based Discriminator

be summarized as follows:
 $\mathbf{f}_t = \sigma(\mathbf{W}_t[\mathbf{h}_{t-1}; \mathbf{k}_t] + \mathbf{b}_t)$ (4) wi 17]. The LSTM C. CNN-based Discrimina

CNN can automatically

widely application in ima

For instance, when process
 $(\mathbf{k}_t + \mathbf{b}_t)$ (5) convolution layer detects e

(6) shapes in the second layer

features like the nos update can be summarized as follows:
 $f_i = \sigma(W_j[h_{i-1}; k_i] + b_j)$ (4) widely application in ima
 $i_i = \sigma(W_j[h_{i-1}; k_i] + b_i)$ (5) convolution layer detects
 $o_i = \sigma(W_k[h_{i-1}; k_i] + b_o)$ (5) convolution layer detects
 $s_i = f_i \, \mathfrak{S}_{i-1} + i$ For instance, when processing an ima
 $\mathbf{v}_i = \sigma(\mathbf{W}_i[\mathbf{h}_{i-1}; \mathbf{k}_i] + \mathbf{b}_i)$ (5) convolution layer detects edges, follow
 $\mathbf{s}_i = \mathbf{f}_i \, \mathbf{S}_{i-1} + \mathbf{i}_i \, \text{C} \mathbf{tanh}(\mathbf{W}_s[\mathbf{h}_{i-1}; \mathbf{k}_i] + \mathbf{b}_i)$ (6) shape $\mathbf{u}_i = \sigma(\mathbf{W}_i | \mathbf{h}_{i-1}; \mathbf{k}_i] + \mathbf{b}_i)$ (5) convolution layer detects edges,
 $\mathbf{o}_i = \sigma(\mathbf{W}_o | \mathbf{h}_{i-1}; \mathbf{k}_i] + \mathbf{b}_o)$ (6) shapes in the second layer, and
 $\mathbf{s}_i = \mathbf{f}_i \times \mathbf{B}_{i-1} + \mathbf{i}_i \times \text{Canh}(\mathbf{W}_s | \mathbf$ These two output vectors are combined to form the final
output of the BiLSTM [18, 19].
The generator includes a BiLSTM layer with 128 neurons,
followed by two Dense layers with 64 and 1 neurons,
respectively. The number o output of the BiLSTM [18, 19].

The generator includes a BiLSTM layer with 128 neurons,

followed by two Dense layers with 64 and 1 neurons,

respectively. The number of neurons in the last layer matches

the output step. The generator includes a BiLSTM layer with 128 neurons,
followed by two Dense layers with 64 and 1 neurons,
respectively. The number of neurons in the last layer matches
the output step. Figure 4 depicts the generation pr followed by two Dense layers with 64 and 1 neurons,
respectively. The number of neurons in the last layer matches
the output step. Figure 4 depicts the generation process of a
single sample input into the generator and de respectively. The number of neurons in the last layer matches
the output step. Figure 4 depicts the generation process of a
single sample input into the generator and defines the overall
output of the generator as \tilde{X} the output step. Figure 4 depicts the generation process of a
single sample input into the generator and defines the overall
output of the generator as $\tilde{X} = (\tilde{X}_4, \tilde{X}_5, ..., \tilde{X}_n)$.
C. CNN-based Discriminator
CNN ca single sample input into the generator and defines the overall
output of the generator as $\tilde{X} = (\tilde{X}_4, \tilde{X}_5, ..., \tilde{X}_n)$.
C. CNN-based Discriminator
CNN can automatically extract deep features and obtain
widely applic output of the generator as $\tilde{X} = (\tilde{X}_4, \tilde{X}_5, ..., \tilde{X}_n)$.
C. CNN-based Discriminator
CNN can automatically extract deep features and obtain
widely application in image classification or text analysis.
For instance, w C. CNN-based Discriminator
CNN can automatically extract deep features and obtain
widely application in image classification or text analysis.
For instance, when processing an image of a dog, the initial
convolution layer C. CNN-based Discriminator
CNN can automatically extract deep features and obtain
widely application in image classification or text analysis.
For instance, when processing an image of a dog, the initial
convolution layer CNN can automatically extract deep features and obtain
widely application in image classification or text analysis.
For instance, when processing an image of a dog, the initial
convolution layer detects edges, followed by widely application in image classification or text analysis.
For instance, when processing an image of a dog, the initial
convolution layer detects edges, followed by the detection of
shapes in the second layer, and the id For instance, when processing an image of a dog, the initial
convolution layer detects edges, followed by the detection of
shapes in the second layer, and the identification of specific
features like the nose in the third convolution layer detects edges, followed by the detection of shapes in the second layer, and the identification of specific features like the nose in the third layer. In the context of time series data, individual data po shapes in the second layer, and the identification of specific
features like the nose in the third layer. In the context of time
series data, individual data points form small trends, which in
turn contribute to larger pat

Volume 51, Issue 8, August 2024, Pages 1105-1111

IAENG International Journal of Computer Science
The discriminator in the proposed IWGAN-GP consists of where $\Pi(P_r, P_s)$ denotes the se
ree one-dimensional (1D) convolutional layers, each with 3, between P_r and P_g , **IAENG International Journal of Computer Science**
The discriminator in the proposed IWGAN-GP consists of where $\Pi(P_r, P_s)$ denotes the set
three one-dimensional (1D) convolutional layers, each with 3, between P_r and P_g **IAENG International Journal of Computer Science**

The discriminator in the proposed IWGAN-GP consists of where $\Pi(P_r, P_g)$ denotes the

three one-dimensional (1D) convolutional layers, each with 3, between P_r and P_g , **IAENG International Journal of Computer Scien**
The discriminator in the proposed IWGAN-GP consists of where $\Pi(P_r, P_g)$ denotes the
three one-dimensional (1D) convolutional layers, each with 3, between P_r and P_g , Π **IAENG International Journal of Computer Science**

The discriminator in the proposed IWGAN-GP consists of where $\Pi(P_r, P_s)$ denotes the stree one-dimensional (1D) convolutional layers, each with 3, between P_r and P_g , **IAENG International Journal of Computer Science**

The discriminator in the proposed IWGAN-GP consists of where $\Pi(P, P_g)$ denotes the s

three one-dimensional (1D) convolutional layers, each with 3, between P_r and P_g , **IAENG International Journal of Computer Science**

The discriminator in the proposed IWGAN-GP consists of where $\Pi(P, P_g)$ denotes the

three one-dimensional (1D) convolutional layers, each with 3, between P_r and P_g , **IAENG Internation**
The discriminator in the proposed IWGAN-GP con
three one-dimensional (1D) convolutional layers, each
2, and 1 convolution kernels respectively, with a unifor
size of 2. To maintain the size of the feat **IAENG International Journal of Computer Science**

mator in the proposed IWGAN-GP consists of where $\Pi(P_i, P_i)$ denotes the set of all jo

isional (1D) convolutional layers, each with 3, between P_r and P_g , Π contain The discriminator in the proposed IWGAN-GP consists of where $\Pi(P, P_g)$ denotes the three one-dimensional (1D) convolutional layers, each with 3, between P_r and P_g , Π con Ω , and 1 convolution kernels respectively The discriminator in the proposed IWGAN-GP consists of where $H(P_r, P_s)$ denotes the s
three one-dimensional (1D) convolutional layers, each with 3, between P_r and P_s , Π conta
2, and I convolution kernels respectivel three one-dimensional (1D) convolutional layers, each with 3, between P_r and P_s , Π cc

2, and 1 convolution kernels respectively, with a uniform step

size of the features during

convolution, the zero-padding meth 2, and 1 convolution kernels respectively, with a uniform s
size of 2. To maintain the size of the features dur
convolution, the zero-padding method is applied. I
one-dimensional data, P rows of zeros are added before
fir **IAENG International Journal of Compute**

proposed IWGAN-GP consists of where $\Pi(P_i, P_{\overline{s}})$ de

convolutional layers, each with 3, between P_r and P_r

respectively, with a uniform step

the size of the features durin and Convolution kernels sepectively, with a uniform steps. As an wind Y , between F_r and F_s . I1 contains ratios of 2. To maintain the size of the features during plans Y . By utilizing the Kantorovolution, the zero-

$$
n_{\text{output}} = (n_{\text{input}} - K + 2 \cdot P) / S + 1 \tag{9}
$$

$$
n_{\text{output}} = \left\lceil \frac{n_{\text{input}}}{S} \right\rceil \tag{10} \begin{array}{c} \text{Compar} \\ \text{the sign} \\ \text{n} \text{rebboli} \end{array}
$$

Contention Convolutional densing Lipschitz contention following Lipschitz contention following Lipschitz contention dentified to the n_{input} *, the WGAN-GP introduces respectively. From Figure 1,* $n_{input} = 4$ *, <i>K* is the convo *n*_{output} = $(n_{input} - K + 2 \cdot P)/3 + 1$ (9) $|f(x_1) - f(x_2)|$

where input and output sizes are denoted by n_{input} and n_{output} , The WGAN-GP introdu

respectively. From Figure 1, $n_{input} = 4$, K is the convolution the Lipschitz constr where input and output sizes are denoted by n_{input} and n_{output} , The WGAN-GP introduces
pectively. From Figure 1, $n_{input} = 4$, K is the convolution the Lipschitz constraint. *L*
kernel, P is the number of filled rows on e respectively. From Figure 1, $n_{input} = 4$, K is the convolution

the Lipschitz constraint. A furthermel, P is the number of filled rows on each side, S is the

starp. In the TensorFlow implementation, the output $1(\|\nab$ respectively. From Figure 1, $n_{input} = 1$, K is the convolution $\frac{1}{K}$

kernel, *P* is the number of filled rows on each side, *S* is the ass

step size. In the TensorFlow implementation, the output 1(

shape is calcula step size. In the TensorFlow implementation, the output $1(\|\nabla f\|_2 \le 1)$ ev

shape is calculated by
 $n_{output} = \begin{bmatrix} \frac{n_{input}}{S} \\ \frac{n_{output}}{S} \end{bmatrix}$ (10) the sigmoid fun

where $\begin{bmatrix} \end{bmatrix}$ represents rounding up. Following th ape is calculated by
 $n_{\text{output}} = \frac{n_{\text{input}}}{S}$ (10) the sigmoid function and ou

erre $\begin{bmatrix} 1 & \text{represents} & \text{rounding up. Following the solution and our probability. This score indicate
\ninvolutional layers, three additional dense layers with 220, represented respectively as
\nd 1 neuron are included. The Rectified Linear Unit (ReLU) $E_{x-k_g}[D(x)] - E_{x-k_g}[D(x)]$ -

represented respectively as

d 1 neuron$ $n_{\text{output}} = \begin{bmatrix} n_{\text{input}} \\ \hline S \end{bmatrix}$ (10) the sigmoid function and outputs

where $\begin{bmatrix} 1 & \text{representes rounding up. Following the data. In the discriminator and ge
convolutional layers, three additional dense layers with 220, represented respectively as
and 1 neuron are included. The Rectified Linear Unit (ReLU) 1
servers as the activation function between these layers, except
for the final layer. Figure 5 is for a visual representation of a
single sample input into the CNN. $\begin{aligned} E_{x \sim P_g}[D(x)] - E_{x \sim P_g}[D(x)] + \lambda E$$ $n_{output} = \frac{|m_{pun}}{S}|$ (10) the sigmoid function are

where $\lceil \rceil$ represents rounding up. Following the data. In the discriminate

convolutional layers, three additional dense layers with 220, represented respectivel!

and where $\lceil \rceil$ represents rounding up. Following the data. In the discriminator and geonvolutional layers, three additional dense layers with 220, represented respectively as and 1 neuron are included. The Rectified Linea where \lceil represents rounding up. Following the data. In the discriminator and g
convolutional layers, three additional dense layers with 220, represented respectively as
and 1 neuron are included. The Rectified Linear

D. The Improved Wasserstein Distance Loss

The loss function of the original GAN is based on KL-JS

introduces two enhancement of the original GAN is based on KL-JS

introduces two enhancement

mimimize the difference b The loss function of the original GAN is based on KL-
divergence. When training, cross-entropy loss is utilized
minimize the difference between the real data distribution
and the generated data distribution, which is equi

$$
-E_{x \sim P_r}[\lg D(x)] - E_{x \sim P_g}[\lg(1 - D(x))]
$$
\n(11)

$$
-E_{x \sim P_{\sigma}}\left[\lg(D(x))\right] \tag{12}
$$

significant issue with JS divergence arises when the two

the gradient penalty section. Firstly, it adopts the L

minimize the difference between the real data distribution

and the generated data distribution, which is e averagner. We can distribution in the gradient penalty section

in the gradient penalty section

minimize the difference between the real data distribution

minimize the KL-JS divergence. The objective function for

the e mannize the Value of the transformation of the discrimination and the generated data distribution, which is equivalent to

minimize the KL-JS divergence. The objective function for

the discriminator is defined as
 $-E_{x\sim$ and the loss function and unionized the the section of the generator is defined as

minimize the KL-JS divergence. The objective function for

the error, the proposed IW

and the loss function for the generator is
 $-E_{x-P_g$ being the error, the proposed IWGAN

the discriminator is defined as
 $-E_{x\sim P_k}[\lg D(x)]-E_{x\sim P_k}[\lg(1-D(x))]$ (11) contains outliers, sacrificing many norm

and the loss function for the generator is
 $-E_{x\sim P_k}[\lg(D(x))]$ (12) additi *Figure 1* and the construments of contains of the set of this information $E_{x-P_k}[B(D(x))]$ and the loss function for the generator is additionally, L₁ regularization $E_{x-P_k}[B(D(x))]$ (12) settings complexity. The discriminator and the loss function for the generator is
 $E_{x-P_k}[\lg(D(x))]$ (11) contains outliers, L₁-ne

and the loss function for the generator is
 $-E_{x-P_k}[\lg(D(x))]$ (12) settings competing to

where E denotes expectation, P_k represen and the loss function for the generator is
 $-E_{x-P_g}[lg(D(x))]$ (12) settings compared to L_2

where E denotes expectation, P_g represents the generated

distribution, and P_r is real data distribution. A

significant issue $-E_{x\text{-}P_x}[lg(D(x))]$ (12) settings compared to

where *E* denotes expectation, P_g represents the generated

data distribution, and P_r is real data distribution. A

significant issue with JS divergence arises when the two
 where E denotes expectation, P_g represents the generated

data distribution, and P_r is real data distribution. A

significant issue with JS divergence arises when the two

distributions have minimal or no overlap, lea IWGAN-GP is:

ignificant issue with JS divergence arises when the two

distributions have minimal or no overlap, leading to JS

distributions have minimal or no overlap, leading to JS

divergence being fixed at a constant significant issue with JS divergence arises when
significant issue with JS divergence arises when
distributions have minimal or no overlap, leadi
divergence being fixed at a constant value of lc
results in the gradient de the inite data distribution is defined as we call that the exist of the second to the second the second to the generator is chemical order of $\epsilon_{x_i, x_j}[B(x)] - E_{x_{i-1}}[B(0-x)]$ (11) outliers, sacrificing many normal samples co **rate distribution**
 rate distribution is delined as the parameter is $-E_{x,x_p}[Eq(D(x))]$ (11) outliers, saerificing many normal sample
 finction for the generator is (11) conditionally, L₁ regularization leads to
 $-E_{x,x_p}[$

$$
W(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} E_{(x, y) \sim \gamma} [\parallel x - y \parallel] \tag{13}
$$

nal of Computer Science
where $\Pi(P_r, P_g)$ denotes the set of all joint distributions
between P_r and P_g , Π contains all the possible transport
plans γ . By utilizing the Kantorovich-Rubinstein duality, the
calcul **nal of Computer Science**
where $\Pi(P_r, P_g)$ denotes the set of all joint distributions
between P_r and P_g , Π contains all the possible transport
plans γ . By utilizing the Kantorovich-Rubinstein duality, the
calcul **nal of Computer Science**

where $\Pi(P_r, P_s)$ denotes the set of all joint distributions

between P_r and P_g , Π contains all the possible transport

plans Y . By utilizing the Kantorovich-Rubinstein duality, the

cal **nal of Computer Science**

where $\Pi(P_r, P_s)$ denotes the set of all joint distributions

between P_r and P_g , Π contains all the possible transport

plans γ . By utilizing the Kantorovich-Rubinstein duality, the

ca Computer Science
 $I(P_r, P_s)$ denotes the set of all joint distributions
 P_r and P_g , Π contains all the possible transport

By utilizing the Kantorovich-Rubinstein duality, the

on can be simplified to
 $(P_r, P_g) = \sup_{\|$ *Let* P_s **,** *E P_s*, **denotes** the set of all joint distributions
 P_s and P_g , Π contains all the possible transport
 Y. By utilizing the Kantorovich-Rubinstein duality, the

ation can be simplified to
 nal of Computer Science

where $\Pi(P_i, P_s)$ denotes the set of all joint distributions

between P_r and P_s , Π contains all the possible transport

plans γ . By utilizing the Kantorovich-Rubinstein duality, the

ca THE THE THE INTERT INTERT IS also the set of all joint distributions
tween P_r and P_g , Π contains all the possible transport
ans γ . By utilizing the Kantorovich-Rubinstein duality, the
lculation can be simplified

$$
W(P_r, P_g) = \sup_{\|f\|_{L} \le 1} E_{x \sim P_r} [f(x)] - E_{x \sim P_g} [f(x)] \qquad (14)
$$

$$
|f(x_1) - f(x_2)| \le |x_1 - x_2|
$$
 (15)

function following Lipschitz constraint as 1 2 1 2 | () () | | | *f x f x x x* (15) where $\Pi(P_1, P_s)$ denotes the set of all joint distributions
between P_r and P_g , Π contains all the possible transport
plans γ . By utilizing the Kantorovich-Rubinstein duality, the
calculation can be simplified t where $\Pi(P_i, P_g)$ denotes the set of all joint distributions
between P_r and P_g , Π contains all the possible transport
plans γ . By utilizing the Kantorovich-Rubinstein duality, the
calculation can be simplified to $1(||\nabla f||_2 \le 1)$ everywhere. If the gradient norm deviates from **If of Computer Science**

nere $\Pi(P, P_s)$ denotes the set of all joint distributions

tween P_r and P_s , Π contains all the possible transport

ans γ' . By utilizing the Kantorovich-Rubinstein duality, the

leulatio plans γ . By utilizing the Kantorovich-Rubinstein duality, the calculation can be simplified to $W(P_r, P_g) = \sup_{\|f\|_{L\leq 1}} E_{x-P_k}[f(x)] - E_{x-P_k}[f(x)]$ (14) where sup is the least upper bound and f is a 1-Lipschitz function fol calculation can be simplified to
 $W(P_r, P_g) = \sup_{\|f\|_{L} \leq 1} E_{x \sim P_r} [f(x)] - E_{x \sim P_g} [f(x)]$ (14)

where sup is the least upper bound and f is a 1-Lipschitz

function following Lipschitz constraint as
 $|f(x_1) - f(x_2)| \leq |x_1 - x_2|$ $W(P_r, P_g) = \sup_{\|f\|_{k} \leq 1} E_{x-P_k}[f(x)] - E_{x-P_k}[f(x)]$ (14)
where sup is the least upper bound and f is a 1-Lipschitz
function following Lipschitz constraint as
 $|f(x_1) - f(x_2)| \leq |x_1 - x_2|$ (15)
The WGAN-GP introduces a gradient pe where sup is the least upper bound and f is a 1-Lipschitz
function following Lipschitz constraint as
 $|f(x_1) - f(x_2)| \le |x_1 - x_2|$ (15)
The WGAN-GP introduces a gradient penalty to enforce
the Lipschitz constraint. A function where sup is the least upper bound and f is a 1-Lipschitz
function following Lipschitz constraint as
 $|f(x_1) - f(x_2)| \le |x_1 - x_2|$ (15)
The WGAN-GP introduces a gradient penalty to enforce
the Lipschitz constraint. A function function following Lipschitz constraint as
 $|f(x_1) - f(x_2)| \le |x_1 - x_2|$ (15)

The WGAN-GP introduces a gradient penalty to enforce

the Lipschitz constraint. A function f is considered

as1-Lipschitz if its gradients have a **Example 12 Example 12** *gre* $\Pi(P_i, P_i)$ denotes the set of all joint distributions
 ween P_r and *P_g*, Π contains all the possible transport
 ns Y. By utilizing the Kantorovich-Rubinstein duality, the

culation can be simplified antorovich-Rubinstein duality, the
 *E*_{x-*P_i*}[*f*(*x*]] - *E_{x-<i>P_i*}[*f*(*x*)] (14)

er bound and *f* is a 1-Lipschitz

tz constraint as
 $(x_2) \le |x_1 - x_2|$ (15)

cas a gradient penalty to enforce

A function *f* is as1-Lipschitz if its gradients have a norm at most
 $x = 10$ and $\| \nabla f \|_2 \le 1$ everywhere. If the gradient norm deviates from

the target norm value of 1, the model will be penalized.

Comparison to the basic GAN, the W 1($\|\nabla f\|_2 \leq 1$) everywhere. If the gradient norm deviates from
the target norm value of 1, the model will be penalized.
Comparison to the basic GAN, the WGAN-GP network lacks
the sigmoid function and outputs a scalar

$$
E_{x \sim P_{g}}[D(x)] - E_{x \sim P_{r}}[D(x)] + \lambda E_{\hat{x} \sim P_{g}}[(\|\nabla_{x}D(x)\|_{2} - 1)^{2}] \quad (16)
$$

$$
-E_{x \sim P_{g}}[D(x)] \quad (17)
$$

keen the multimate of the lead tower and state. Since the energy and the base calculated by

step size. In the TensorFlow implementation, the output $1(|Vf|)|_2 \le 1$) everywhere. If the gradient is

shape is calculated by
 y. From Figure 1, $n_{\text{gamma}} = 4$, *K* is the convolution the Lipschitz of its gradients.

is the number of filled rows on each side, *S* is the and -1-lipschitz if its gradients have

In the TensorFlow implementation, the and the big means of the signal control in the sign of the signal completion for the basic GAN, it

where $\left[\begin{array}{c} \frac{1}{n_{\text{input}}} \end{array}\right]$ represents rounding up. Following the data. In the discriminator and usure,

and 1 n *g E D x x P* (12) Fustance Eoss

incroduces two enhancements to

inconsecution. Firstly, it adopts the L₁

oss-entropy loss is utilized to

een the real data distribution

time different predict proposed in the gradient penalty section
 For the target norm value of 1, the model will be penalized.
Comparison to the basic GAN, the WGAN-GP network lacks
the sigmoid function and outputs a scalar score rather than a
probability. This score indicates the authe the target norm value of 1, the model will be penalized.
Comparison to the basic GAN, the WGAN-GP network lacks
the sigmoid function and outputs a scalar score rather than a
probability. This score indicates the authentic Comparison to the basic OATN, the WOATN-OF hetwork lacks
the sigmoid function and outputs a scalar score rather than a
probability. This score indicates the authenticity of the input
data. In the discriminator and generat the sigmola function and outputs a scalar scole rather than a
probability. This score indicates the authenticity of the input
data. In the discriminator and generator, the loss function is
represented respectively as
 E_{x probability. This scole indicates the authenticity of the input
data. In the discriminator and generator, the loss function is
represented respectively as
 $E_{x\rightarrow P_g}[D(x)] + \lambda E_{x\rightarrow P_g}[(\|\nabla_x D(x)\|_2 - 1)^2]$ (16)
 $-E_{x\rightarrow P_g}[D(x)]$ (17 data. In the discriminator and generator, the loss function is
represented respectively as
 $E_{x\sim P_g}[D(x)] - E_{x\sim P_g}[D(x)] + \lambda E_{x\sim P_g}[(\|\nabla_x D(x)\|_2 - 1)^2]$ (16)
 $-E_{x\sim P_g}[D(x)]$ (17)
where \hat{x} represents the intermediate value bet represented respectively as
 $E_{x-P_g}[D(x)] - E_{x-P_g}[D(x)] + \lambda E_{\hat{x}-P_g}[(\|\nabla_x D(x)\|_2 - 1)^2]$ (16)
 $-E_{x-P_g}[D(x)]$ (17)

where \hat{x} represents the intermediate value between the real

and generated sample space. The proposed IWGAN-GP
 $E_{x\sim P_g}[D(x)] - E_{x\sim P_g}[D(x)] + \lambda E_{\hat{x}\sim P_g}[(\|V_xD(x)\|_2 - 1)^2]$ (16)
 $-E_{x\sim P_g}[D(x)]$ (17)

where \hat{x} represents the intermediate value between the real

and generated sample space. The proposed IWGAN-GP

introduces two enhancem $-E_{x-P_g}[D(x)]$ (17)
where \hat{x} represents the intermediate value between the real
and generated sample space. The proposed IWGAN-GP
introduces two enhancements to the basic WGAN-GP loss
function. Firstly, it adopts the L₁ where \hat{x} represents the intermediate value between the real
and generated sample space. The proposed IWGAN-GP
introduces two enhancements to the basic WGAN-GP loss
function. Firstly, it adopts the L₁ norm instead of where *x* represents the intermediate value between the real
and generated sample space. The proposed IWGAN-GP
introduces two enhancements to the basic WGAN-GP loss
function. Firstly, it adopts the L₁ norm instead of th and generated sample space. The proposed IW
introduces two enhancements to the basic WGAN
function. Firstly, it adopts the L₁ norm instead of the
in the gradient penalty section of the discrimin
function. The L₁ norm discriminator and generator, the loss function is
discriminator and generator, the loss function is
respectively as
 x)] $-E_{x-P_z}[D(x)] + \lambda E_{x-P_z}[(||\nabla_x D(x)||_2 - 1)^2]$ (16)
 $-E_{x-P_z}[D(x)]$ (17)
presents the intermediate value between *y*. This score that outgous a scalar scote radio and the during the system in the discriminator and generator, the loss function is despectively as (xx) $E_{x-P_z}[D(x)] + \lambda E_{x-P_z}[[U(x)]$ (17) (17) $E_{x-P_z}[D(x)] + \lambda E_{x-P_z}[U(x)]$ (17) (1 Lettion. Firstly, it adopts the L₁ horm instead of the L₂ horm
the gradient penalty section of the discriminator loss
nction. The L₁ norm exhibits better tolerance to outliers
mpared to the L₂ norm. This is becaus in the gradient penarty section of the discriminator loss
function. The L₁ norm exhibits better tolerance to outliers
compared to the L₂ norm. This is because the L₂ norm squares
the error, the proposed IWGAN-GP wil function. The L₁ norm extinuts better toterance to outners
compared to the L₂ norm. This is because the L₂ norm squares
the error, the proposed IWGAN-GP will be more sensitive to
outliers, sacrificing many normal sa compared to the L₂ norm. This is because the L₂ norm squares
the error, the proposed IWGAN-GP will be more sensitive to
outliers, sacrificing many normal samples. When the dataset
contains outliers, L₁-norm is more **EXECUTE TO EXECUTE THE CONDUCT CONDUCTS**
 EXECUTE ANOTE AND CONDUCTS (**C**) $\log(AN - G)$ and $\log(AN - G)$ will be more sensitive to many normal samples. When the dataset

and IWGAN-GP will be more sensitive to many normal sam may be space. Ine proposed IWGAN-GP
Inhamements to the basic WGAN-GP
Inhamements to the basic WGAN-GP loss
it adopts the L₁ norm instead of the L₂ norm
penalty section of the discriminator loss
norm exhibits better to

$$
E_{x \sim P_g} [D(x)] - E_{x \sim P_r} [D(x)] + \lambda E_{\hat{x} \sim P_s} [(\|\nabla_x D(x)\|_1 - 1)^2] \tag{18}
$$

the error, the proposed IWGAN-GP will be flore sensitive to outliers, sacrificing many normal samples. When the dataset
contains outliers, L₁-norm is more effective than L₂-norm.
Additionally, L₁ regularization lead butters, sacrincing many normal samples. When the diaset

contains outliers, L₁-norm is more effective than L₂-norm.

Additionally, L₁ regularization leads to sparser parameter

settings compared to L₂ regularizat by: Implexity. The discriminator loss function in the proposed
 $V(\text{GAN-GP is:}\nE_{x\text{-}P_k}[D(x)] - E_{x\text{-}P_k}[D(x)] + \lambda E_{x\text{-}P_k}[(\|\nabla_x D(x)\|_1 - 1)^2]$ (18)

Secondly, the Mean Squared Error (MSE) between real

mples and the generated sampl IWGAN-GP is:
 $E_{x\text{-}P_k}[D(x)] - E_{x\text{-}P_k}[D(x)] + \lambda E_{x\text{-}P_k}[(\|\nabla_x D(x)\|_1 - 1)^2]$ (18)

Secondly, the Mean Squared Error (MSE) between real

samples and the generated samples is incorporated into the

generator's loss function. E_{x-P_s}[D(x)]–E_{x-P_s}[D(x)]+ $\lambda E_{\hat{x}-P_{\hat{x}}}$ [(|| $\nabla_x D(x)$ |
Secondly, the Mean Squared Error (MSE) b
samples and the generated samples is incorpor-
generator's loss function. This addition aims to
model stability by ad can Squared Error (MSE) between real
erated samples is incorporated into the
tion. This addition aims to improve the
ljusting the generator based on the MSE
ator makes incorrect judgments. The
tion for enhancing WGAN-GP i generator's loss function. This addition aims to impromodel stability by adjusting the generator based on the
when the discriminator makes incorrect judgments
generator's loss function for enhancing WGAN-GP is
by:
 $- E_{x \sim$

$$
-E_{x \sim P_g}[D(x)] - E[(x_{P_r} - x_{P_g})^2]
$$
\n(19)

ball stability by adjusting the generator based on the MSE

nen the discriminator makes incorrect judgments. The

nerator's loss function for enhancing WGAN-GP is given

:
 $- E_{x \cdot P_k}[D(x)] - E[(x_p - x_{P_k})^2]$ (19)

Table I compa when the discriminator makes incorrect judgments. The
generator's loss function for enhancing WGAN-GP is given
by:
 $- E_{s-h} [D(x)] - E[(x_h - x_h)^2]$ (19)
Table I compares the loss functions of generators and
discriminators in the b generator's loss function for enhancing WGAN-GP is given
by:

 $-E_{s-h_i}[D(x)]-E[(x_h-x_{h_i})^2]$ (19)
Table I compares the loss functions of generators and
discriminators in the basic GAN, basic WGAN-GP, and
proposed IWGAN-GP.
III. by:
 $-E_{s-t_k}[D(x)]-E[(x_{t_k}-x_{t_k})^2]$ (19)

Table I compares the loss functions of generators and

discriminators in the basic GAN, basic WGAN-GP, and

proposed IWGAN-GP.

III. EXPERIMENTS

A. Evaluation index

Similar to the $-E_{x\text{-}P_x}[D(x)] - E[(x_{P_x} - x_{P_x})^2]$ (19)

Table I compares the loss functions of generators and

discriminators in the basic GAN, basic WGAN-GP, and

proposed IWGAN-GP.

III. EXPERIMENTS

A. Evaluation index

Similar to the Table I compares the loss functions of generators and
scriminators in the basic GAN, basic WGAN-GP, and
poposed IWGAN-GP.
III. EXPERIMENTS
Evaluation index
Similar to the classical evaluating metrics of regression
gorithms

IAENG International Journal of Computer Science
 $\widetilde{X} = (\widetilde{x}_4, \widetilde{x}_5, ..., \widetilde{x}_n)$ is denormalized to obtain the predicted

LOSS FUNCTIONS OF THE DISCRIM

values $Y = (y_4, y_5, ..., y_n)$; given the original data
 $X = (x_4, x_5$ **IAENG International Journal of Compu**
 $\widetilde{X} = (\widetilde{x}_4, \widetilde{x}_5, ..., \widetilde{x}_n)$ is denormalized to obtain the predicted

values $Y = (y_4, y_5, ..., y_n)$; given the original data
 $X = (x_4, x_5, ..., x_n)$. These metrics are calculated as f **IAENG International Journal of Computer Sci**
is denormalized to obtain the predicted
..., y_n) ; given the original data $\underbrace{S_{\text{Model}}}_{\text{Model}}$ Discriminal characterics are calculated as follows
Frror (MSF): A commonly us **IAENG International Journal of Computer S**
 $\widetilde{X} = (\widetilde{x}_4, \widetilde{x}_5, ..., \widetilde{x}_n)$ is denormalized to obtain the predicted

values $Y = (y_4, y_5, ..., y_n)$; given the original data DIFFE
 $X = (x_4, x_5, ..., x_n)$. These metrics are calcu **IAENG International Journal of Computer Science**
 $\widetilde{X} = (\widetilde{x}_4, \widetilde{x}_5, ..., \widetilde{x}_n)$ is denormalized to obtain the predicted

values $Y = (y_4, y_5, ..., y_n)$; given the original data **DIFFERENT G**

values $X = (x_4, x_5, ..., x_n)$. Th **IAENG International Journal of Computer Scion**

1) is denormalized to obtain the predicted
 $y_5,..., y_n$; given the original data <u>DIFFERE</u>

These metrics are calculated as follows

These metrics are calculated as follows
 14ENG International Journal of Computer Scien
 $I = (\tilde{x}_4, \tilde{x}_5, ..., \tilde{x}_n)$ is denormalized to obtain the predicted

lues $Y = (y_4, y_5, ..., y_n)$; given the original data *DIFFERENT*
 $I = (x_4, x_5, ..., x_n)$. These metrics are calcul **IAENG International Journal of Computer Science (A)**
 $\widetilde{X} = (\widetilde{x}_4, \widetilde{x}_5, ..., \widetilde{x}_n)$ is denormalized to obtain the predicted

values $Y = (y_4, y_5, ..., y_n)$; given the original data <u>DISS FUNCTIONS</u> OF THE DIS
 $X = (x_4, x_$ **IAENG International Journal of Computer**:
 $\widetilde{X} = (\widetilde{x}_4, \widetilde{x}_5, ..., \widetilde{x}_n)$ is denormalized to obtain the predicted

values $Y = (y_4, y_5, ..., y_n)$; given the original data $\frac{\text{Model}}{\text{Model}}$
 $X = (x_4, x_5, ..., x_n)$. These metrics a

$$
MSE = \frac{1}{n-3} \sum_{t=4}^{n} (y_t - x_t)^2
$$
 (20)

$$
RMSE = \sqrt{\frac{1}{n-3} \sum_{t=4}^{n} (y_t - x_t)^2}
$$
 result
(21) IWC
follows follow

performance. *B.* Settings

In evaluating the performance.
 B. Settings
 B. Settings

In evaluating the proposed IWAGN-GP

basic GAN and the basic WGAN-C

basic GAN and the basic WGAN-C

basic GAN and the basic WGAN-C

basic Herma

$$
MAE = \frac{1}{n-3} \sum_{t=4}^{n} |y_t - x_t|
$$
 (22) 41.12%, 47.17%, and
WGAN-GP.

$$
MAPE = \frac{1}{n-3} \sum_{t=4}^{n} \frac{|y_t - x_t|}{x_t}
$$
 (23)
BilS
short. (23)

 $MAE = \frac{1}{n-3} \sum_{t=4}^{n} |y_t - x_t|$ (22) WGAN-GP.

4) Mean Absolute Percentage Error (MAPE): The Table IV displays the paverage of the percentage differences between predicted data oftained by different methods results indica 4) Mean Absolute Percentage Error (MAPE): The Table IV displays
average of the percentage differences between predicted data
and original data, often used to measure relative error size.
Table IV, the prop
and original da 4) Mean Absolute Percentage Error (MAPE): The
average of the percentage differences between predicted data
and original data, often used to measure relative error size.
and original data, often used to measure relative er average of the percentage differences between predicted data

and original data, often used to measure relative error size.

Table IV, the proposed
 $MAPE = \frac{1}{n-3} \sum_{t=4}^{n} \frac{|y_t - x_t|}{x_t}$ (23) BiLSTM and CNN, effective

Bi and original data, often used to measure relative error size.
 $MAPE = \frac{1}{n-3} \sum_{i=4}^{n} \frac{|y_i - x_i|}{x_i}$ (23) BiLSTM and CNN, effectively

Short-term characteristics of

B. Settings

In evaluating the proposed IWAGN-GP, the B $MAPE = \frac{1}{n-3} \sum_{t=4}^{n} \frac{|y_t - x_t|}{x_t}$ (23) BiLSTM and CNN, effect

B. Settings

B. Settings

B. Settings

D. Settings

B. Settings

D. Settings

D. Settings

D. Settings

D. Setting in improved fitting the proposed IWAGN*n* $nH = \frac{1}{2}$ $n = 3$ and $n = 4$ and $n = 5$ and $n = 6$ and $n = 6$ *B. Settings*

In evaluating the proposed IWAGN-GP, the BiLSTM, the

the relative errors calculated fro

basic GAN and the basic WGAN-GP are also be

implemented for comparison, of which Both the GAN and

the WGAN-GP adopt B. Settings

In evaluating the proposed IWAGN-GP, the BiLSTM, the

basic GAN and the basic WGAN-GP are also be

implemented for comparison, of which Both the GAN and

implemented for comparison, of which Both the GAN and
 In evaluating the proposed *WAGN-GH*
basic GAN and the basic WGAN-C
implemented for comparison, of which B
the WGAN-GP adopt the BiLSTM as the g
as the discriminator. All models utilize Ad
algorithm with learning rates sel Fraction results of BiLSTM, GAN, WGAN-GP, and

Fraction be seen that the predicted at the predicted at the prediction results of the bise

plemented for comparison, of which Both the GAN and

TWGAN-GP for the basic C

the INCORN-GP are depicted in Figure 6. Figure 6. INCORN-GP L₁ norm instead of the L₂ norm

Figure 6 (iv) complex that the proposed IWGAN-GP over the basic departments, it is found that lower

algorithm with learning rate and the discriminator. All models utilize Adam's optimization
algorithm with learning rates selected from 0.0001, 0.0003, 3.70%, respectively. IWGA
0.001, to 0.003. Through experiments, it is found that lower
algorithm wi

algorithm with learning rates selected from 0.0001, 0.0003,

0.001, to 0.003. Through experiments, it is found that lower

0.001, to 0.003. Through experiments, it is found that lower

learning rates lead to smoother loss (Figure 6 (d)) demonstrates a better fit to actual data better fit demonstrates a better fit to actual data better fit to actual data better fi Example the state of the station results of BiLSTM, GAN and WGAN-GP, respectively improved prediction results. Therefore, the optimal learning
improved prediction results. Therefore, the optimal learning
the stability of Example the production results. Therefore, the optimal learning

improved prediction results. Therefore, the optimal learning

the stability of the loss function curve, and the random

the stability of the loss function c The issue to 0.0001, the number of cycles is determined by
the stability of the loss function curve, and the random
number seed for IWGAN-GP is set to 5.

C. Results

C. Results

C. Results

C. Results

C. Results

C. Res depicted in Figure 2. Comparing Figure 6 (b) and Figure 6 (c), and be prediction research in Figure 6. Figure 6 (d) demonstrates a better fit to actual data trend. The prediction of a new WGAN-GP is set to 5.

C. Results
 number seed for IWGAN-GP is set to 5.

C. *Results*

The prediction results of BiLSTM, GAN, WGAN-GP, and

IWGAN-GP are depicted in Figure 6. Figure 6 illustrates that

IWGAN-GP are depicted in Figure 6. Figure 6 illustrate 1) construction of a new WG

C. Results

The prediction results of BiLSTM, GAN, WGAN-GP, and

The prediction MSE in the series

IWGAN-GP are depicted in Figure 6. Figure 6 illustrates that

difference between generated

a *C. Results*
 Example 10 The prediction results of BiLSTM, GAN, WGAN-GP, and extraction of time series fe

IWGAN-GP are depicted in Figure 6. Figure 6 illustrates that

the difference between generator

and the seen tha The prediction results of BiLSTM, GAN, WGAN-GP, and

extractic

IWGAN-GP are depicted in Figure 6. Figure 6 illustrates that

predictit

it can be seen that the predicted data trend aligns with the

differen

actual data t GAN-GP are depicted in Figure 6. Figure 6 illustrates that

trace the seen that the predicted data trend aligns with the

trace between generated and

tual data trend. However, the proposed IWGAN-GP L₁ norm instead of t it can be seen that the predicted data trend aligns with the difference between generated actual data trend. However, the proposed IWGAN-GP L_1 norm instead of the L_2 nor (Figure 6 (d)) demonstrates a better fit to a actual data trend. However, the proposed IWGAN-GP L₁ norm instead of the L₂ norm i (Figure 6 (d)) demonstrates a better fit to actual data the discriminator loss function to compared to the other three models. In Figur (Figure 6 (d)) demonstrates a better fit to actual data the discriminator loss functio compared to the other three models. In Figure 6 (a), the robustness. It is verified by BiLSTM prediction results show a nearly straigh

nce

TABLE I

SCRIMINATOR AND GENERATOR OF

NT GAN MODELS

Generator

IAENG International Journal of Computer Science				
$\widetilde{X} = (\widetilde{x}_4, \widetilde{x}_5, , \widetilde{x}_n)$ is denormalized to obtain the predicted	TABLE I LOSS FUNCTIONS OF THE DISCRIMINATOR AND GENERATOR OF DIFFERENT GAN MODELS			
values $Y = (y_4, y_5, , y_n)$; given the original data	Model	Discriminator	Generator	
$X = (x_4, x_5, , x_n)$. These metrics are calculated as follows 1) Mean Squared Error (MSE): A commonly used metric	basic GAN	$-E_{x\sim P}$ [lg D(x)] $-E_{x-P_{x}}[lg(1-D(x))]$	$-E_{x\sim P_{\sigma}}[lg(D(x))]$	
in regression analysis, defined as the average of the squared differences between predicted data and original data. $MSE = \frac{1}{n-3}\sum_{i=1}^{n} (y_i - x_i)^2$ (20)	basic WGAN-GP	$E_{x \sim P_n}[D(x)] - E_{x \sim P_n}[D(x)]$ $+ \lambda E_{\hat{x} \sim P} [(\nabla_x D(x) _2 - 1)^2]$	$-E_{_{x\sim P_{\sigma}}}[D(x)]$	
		$E_{x \sim P_n}[D(x)] - E_{x \sim P_n}[D(x)]$	$-E_{x\sim P_{x}}[D(x)]$	
2) Root Mean Squared Error (RMSE): The square root of the average of the squares of prediction errors, which considers the magnitude of the prediction errors and	Proposed IWGAN-GP	$+ \lambda E_{\hat{x} \sim P} [(\nabla_x D(x) _1 - 1)^2]$	$-E[(x_{p} - x_{p})^{2}]$	
penalizes larger errors more. $RMSE = \sqrt{\frac{1}{n-3}\sum_{i=1}^{n}(y_i - x_i)^2}$ (21) 3) Mean Absolute Error (MAE): The average of the absolute differences between predicted data and original data. The smaller MAE represents the better predictive performance. $MAE = \frac{1}{n-3}\sum_{i=1}^{n} y_i - x_i $ (22)	WGAN-GP.	ing datasets and Table III for test datasets, with the best results highlighted in bold. The analysis reveals that IWGAN-GP shows the best predictive performance, followed by WGAN-GP, while GAN performs the weakest. A comparison of the four metrics shows that IWGAN-GP demonstrates lower predictive errors, reducing MSE, RMSE, MAE, and MAPE on the test set by 72.33%, 70.87%, 72.72%, and 46.62% respectively compared to GAN; and by 66.33%, 41.12%, 47.17%, and 41.34% respectively compared to		
4) Mean Absolute Percentage Error (MAPE): The average of the percentage differences between predicted data and original data, often used to measure relative error size. $MAPE = \frac{1}{n-3}\sum_{t=4}^{n} \frac{ y_t - x_t }{x_t}$ (23)		Table IV displays the grain consumption forecasts obtained by different methods from 2016 to 2020. The bolded results indicate the values closest to the actual values. From Table IV, the proposed IWGAN-GP, which includes BiLSTM and CNN, effectively captures both long-term and		

 $=\sqrt{\frac{1}{n-3}\sum_{t=4}^{n}(y_t-x_t)^2}$ (21) IWGAN-GP shows the best predictive performs the vector followed by WGAN-GP while GAN performs the w 1) Mean Squared Error (MSE): A commonly used metric

regression analysis, defined as the average of the squared

Terences between predicted data and original data.
 $MSE = \frac{1}{n-3} \sum_{i=4}^{n} (y_i - x_i)^2$ (20)

2) Root Mean Squa in regression analysis, defined as the average of the squared

differences between predicted data and original data.
 $MSE = \frac{1}{n-3} \sum_{i=4}^{n} (y_i - x_i)^2$ (20)

2) Root Mean Squared Error (RMSE): The square root of Proposed
 differences between predicted data and original data. basic $\mu_{x} = \frac{1}{n-3} \sum_{i=4}^{n} (y_i - x_i)^2$ (20)

2) Root Mean Squared Error (RMSE): The square root of $\mu_{x} = \mu_{x} [D(x_0 + x_0)]$

2) Root Mean Squared Error (RMSE): The s **EXECUTE:** The square root of Proposed
diction errors, which IWGAN-GP
prediction errors and $\frac{1}{\log \text{ datasets}}$
 $\frac{1}{1-\frac{x_i}{2}}$ (21) IWGAN-GP
followed by N
i. The average of the A comparison
data and original data. demonstrate 41.12%, 47.17%, and 41.34% respectively compared to
WGAN-GP.
Table IV displays the grain consumption forecasts 2) average of the squares of prediction errors, which

malizes larger errors more.

malizes larger errors more.

malizes larger errors more.

MASE = $\sqrt{\frac{1}{n-3} \sum_{t=4}^{n} (y_t - x_t)^2}$ (21) IWGAN-GP shows the beat

3) Mean A considers the magnitude of the prediction errors and

penalizes larger errors more.
 $RMSE = \sqrt{\frac{1}{n-3} \sum_{t=4}^{n} (y_t - x_t)^2}$ (21) IWGAN-GP shows the bs

3) Mean Absolute Error (MAE): The average of the A comparison of the fou basic $E_{x\rightarrow P_k}[g(1-D(x))]$

basic $E_{x\rightarrow P_k}[D(x)] - E_{x\rightarrow P_k}[D(x)]$

WGAN-GP $+ \lambda E_{\hat{x}\rightarrow P_k}[(\|\nabla_x D(x)\|_2 - 1)^2]$ $- E_{x\rightarrow P_k}[D(x)]$

Proposed $E_{x\rightarrow P_k}[D(x)] - E_{x\rightarrow P_k}[D(x)]$

IWGAN-GP $+ \lambda E_{\hat{x}\rightarrow P_k}[(\|\nabla_x D(x)\|_1 - 1)^2]$ $- E[(x_p - x_p)^2]$

ing datase basic $E_{s\text{-}P_k}[D(x)] - E_{s\text{-}P_k}[D(x)]$

WGAN-GP $+ \lambda E_{\hat{s}\text{-}P_k}[D(x)||_2 - 1)^2]$ $- E_{s\text{-}P_k}[D(x)]$

Proposed $E_{s\text{-}P_k}[D(x)] - E_{s\text{-}P_k}[D(x)]$

IWGAN-GP $+ \lambda E_{\hat{s}\text{-}P_k}[(|\nabla_x D(x)||_1 - 1)^2]$ $- E[(x_p - x_{p_k})^2]$

ing datasets and Table III basic

WGAN-GP $+ \lambda E_{\frac{r}{2}, P_{\frac{r}{2}}}[(||\nabla_x D(x)||_2 - 1)^2]$ $- E_{\frac{r}{2}, P_{\frac{r}{2}}}[D(x)]$

Proposed $E_{\frac{r}{2}, P_{\frac{r}{2}}}[(||\nabla_x D(x)||_1 - 1)^2]$ $- E[(x_p - x_{P_{\frac{r}{2}}})^2]$

IWGAN-GP $+ \lambda E_{\frac{r}{2}, P_{\frac{r}{2}}}[(||\nabla_x D(x)||_1 - 1)^2]$ $- E[(x_p - x_{P_{\frac$ WGAN-GP $H.E_{x,P_k}[(\|V_xD(x)\|_2 - 1)^2]$ $- E_{x-P_k}[D(x)]$

Proposed $E_{x-P_k}[D(x) - 1]^2$ $- E_{x-P_k}[D(x)]$

IWGAN-GP $H.E_{x,P_k}[(\|V_xD(x)\|_1 - 1)^2]$ $- E[(x_p - x_p)^2]$

ing datasets and Table III for test datasets, with the best

results highlighted Proposed $E_{x\sim P_k}[D(x)] - E_{x\sim P_k}[D(x)]$ $-E_{x\sim P_k}[D(x)]$

IWGAN-GP $+ \lambda E_{x\sim P_k}[(\|\nabla_x D(x)\|_1 - 1)^2]$ $-E[(x_p - x_{P_k})^2]$

ing datasets and Table III for test datasets, with the best

results highlighted in bold. The analysis reveals t VGAN-GP $+ \lambda E_{x,P_s}[(\|\nabla_x D(x)\|_1 - 1)^2]$ $- E[(x_{P_s} - x_{P_s})^2]$

g datasets and Table III for test datasets, with the best

sults highlighted in bold. The analysis reveals that

VGAN-GP shows the best predictive performance,
 ing datasets and Table III for test datasets, with the best
results highlighted in bold. The analysis reveals that
IWGAN-GP shows the best predictive performance,
followed by WGAN-GP, while GAN performs the weakest.
A comp ing datasets and Table III for test datasets, with the best
results highlighted in bold. The analysis reveals that
IWGAN-GP shows the best predictive performance,
followed by WGAN-GP, while GAN performs the weakest.
A comp results highlighted in bold. The analysis reveals that IWGAN-GP shows the best predictive performance, followed by WGAN-GP, while GAN performs the weakest. A comparison of the four metrics shows that IWGAN-GP demonstrates

differences between predicted data and original data.
 $MSE = \frac{1}{n-3} \sum_{i=4}^{n} (y_i - x_i)^2$ (20)

20) WGAN-GP are also be grain considers the manging of the sugness of proposed $E_{x_i}f(0|\nabla_iD(x)|_i-1)^2$

2) Root Mean Squared E *th* (21) IWONN of followed by

(21) IWONN of followed by

(21) IWONN of the step of the A comparise

d data and original data. demonstrate:

the better predictive MAE, and M

and 46.62%
 $y_t - x_t$ (22) $\begin{array}{c} 41.12\% , 47$ een predicted data and original data.
 MSE = $\frac{1}{n-3} \sum_{i=4}^{n} |y_i - x_i|$ (20)
 S Squared Error (RMSE): The square root of
 Formoscolar *Formoscolar Fig.* $p(x) = E_{x,x_E}[U(\nabla_x D(x) - E_{x,x_E}[D(x)] - E_{y,E}[D(x)]$

the squares of pre $=\frac{1}{n-3}\sum_{i=4}^{n} (y_i - x_i)^2$ (20)

Error (RMSE): The square root of Proposed $F_{x \cdot p_i}[U|\nabla_z D(x)|]_+$

Error (RMSE): The square root of Proposed $F_{x \cdot p_i}[U|\nabla_z D(x)|]_+$

e. of the prediction errors, which IWGAN-GP $+ \lambda E_{x \cdot p$ Table IV, the proposed IWGAN-GP, which includes
BiLSTM and CNN, effectively captures both long-term and
short-term characteristics of grain consumption trends. Example 18 MAE represents the better predictive MAE, and MAPE on the test set

formance.

In all 46.62% respectively compared and 46.62% respectively compared in the test set

of the proposed ifferences between pred berformance.

The manner of the precisents and 46.62% respectively compare

and 46.62% respectively compare

4) Mean Absolute Percentage Error (MAPE): The Table IV displays the grad

average of the percentage differences MAE = $\frac{1}{n-3} \sum_{i=4}^{n} |y_i - x_i|$ (22) $\frac{41.12\%}{WGAN-GP}$, and 41.349

4) Mean Absolute Percentage Error (MAPE): The Table IV displays the g

average of the percentage differences between predicted data

and original dat IWGAN-GP shows the best predictive performance,
followed by WGAN-GP, while GAN performs the weakest.
A comparison of the four metrics shows that IWGAN-GP
demonstrates lower predictive errors, reducing MSE, RMSE,
MAE, and M followed by WGAN-GP, while GAN performs the weakest.
A comparison of the four metrics shows that IWGAN-GP
demonstrates lower predictive errors, reducing MSE, RMSE,
MAE, and MAPE on the test set by 72.33%, 70.87%, 72.72%,
a A comparison of the four metrics shows that IWGAN-GP
demonstrates lower predictive errors, reducing MSE, RMSE,
MAE, and MAPE on the test set by 72.33%, 70.87%, 72.72%,
and 46.62% respectively compared to GAN; and by 66.33% demonstrates lower predictive errors, reducing MSE, RMSE,
MAE, and MAPE on the test set by 72.33%, 70.87%, 72.72%,
and 46.62% respectively compared to GAN; and by 66.33%,
41.12%, 47.17%, and 41.34% respectively compared to MAE, and MAPE on the test set by 72.33%, 70.87%, 72.72%,
and 46.62% respectively compared to GAN; and by 66.33%,
41.12%, 47.17%, and 41.34% respectively compared to
WGAN-GP.
Table IV displays the grain consumption forecast and 46.62% respectively compared to GAN; and by 66.33%,
41.12%, 47.17%, and 41.34% respectively compared to
WGAN-GP.
Table IV displays the grain consumption forecasts
obtained by different methods from 2016 to 2020. The bo 41.12%, 47.17%, and 41.34% respectively compared to
WGAN-GP.
Table IV displays the grain consumption forecasts
obtained by different methods from 2016 to 2020. The bolded
results indicate the values closest to the actual v WGAN-GP.
Table IV displays the grain consumption forecasts
obtained by different methods from 2016 to 2020. The bolded
results indicate the values closest to the actual values. From
Table IV, the proposed IWGAN-GP, which i Table IV displays the grain consumption forecasts
obtained by different methods from 2016 to 2020. The bolded
results indicate the values closest to the actual values. From
Table IV, the proposed IWGAN-GP, which includes
B obtained by different methods from 2016 to 2020. The bolded
results indicate the values closest to the actual values. From
Table IV, the proposed IWGAN-GP, which includes
BiLSTM and CNN, effectively captures both long-term results indicate the values closest to the actual values. From
Table IV, the proposed IWGAN-GP, which includes
BiLSTM and CNN, effectively captures both long-term and
short-term characteristics of grain consumption trends, effectively captures both long-term and
istics of grain consumption trends,
I fitting performance. Table V presents
lculated from the actual and predicted
from 2016 to 2020 using different
mallest errors highlighted in bol ort-term characteristics of grain consumption trends,
sulting in improved fitting performance. Table V presents
relative errors calculated from the actual and predicted
ain consumption from 2016 to 2020 using different
tho resulting in improved fitting performance. Table V presents
the relative errors calculated from the actual and predicted
grain consumption from 2016 to 2020 using different
methods, with the smallest errors highlighted in the relative errors calculated from the actual and predicted
grain consumption from 2016 to 2020 using different
methods, with the smallest errors highlighted in bold. The
average errors for the basic GAN, basic WGAN-GP, a

gram consumption from 2016 to 2020 using difterent
methods, with the smallest errors highlighted in bold. The
average errors for the basic GAN, basic WGAN-GP, and
IWGAN-GP over these five years are 7.40%, 7.95%, and
3.70%, methods, with the smallest errors highlighted in bold. The
average errors for the basic GAN, basic WGAN-GP, and
IWGAN-GP over these five years are 7.40%, 7.95%, and
3.70%, respectively. IWGAN-GP has increased the
predicti average errors for the basic GAN, basic WGAN-GP, and
IWGAN-GP over these five years are 7.40%, 7.95%, and
3.70%, respectively. IWGAN-GP has increased the
prediction accuracy by 50% and 53.46% compared to the
GAN and WGAN-IWGAN-GP over these tive years are 7.40%, 7.95%, and
3.70%, respectively. IWGAN-GP has increased the
prediction accuracy by 50% and 53.46% compared to the
GAN and WGAN-GP, respectively.
IV. CONCLUSION
A new WGAN-GP model 3.70%, respectively. IWGAN-GP has increased the
prediction accuracy by 50% and 53.46% compared to the
GAN and WGAN-GP, respectively.
IV. CONCLUSION
A new WGAN-GP model is proposed for grain
consumption prediction, involvi prediction accuracy by 50% and 53.46% compared to the
GAN and WGAN-GP, respectively.
IV. CONCLUSION
A new WGAN-GP model is proposed for grain
consumption prediction, involving the following key aspects:
1) construction of GAN and WGAN-GP, respectively.

IV. CONCLUSION

A new WGAN-GP model is proposed for grain

consumption prediction, involving the following key aspects:

1) construction of a new WGAN-GP using BiLSTM as the

generator and IV. CONCLUSION

A new WGAN-GP model is proposed for grain

consumption prediction, involving the following key aspects:

1) construction of a new WGAN-GP using BiLSTM as the

generator and CNN as the discriminator to enha IV. CONCLUSION
A new WGAN-GP model is proposed for grain
consumption prediction, involving the following key aspects:
1) construction of a new WGAN-GP using BiLSTM as the
generator and CNN as the discriminator to enhance A new WGAN-GP model is proposed for grain consumption prediction, involving the following key aspects: 1) construction of a new WGAN-GP using BiLSTM as the generator and CNN as the discriminator to enhance the extraction

Year	Actual	Predicted Values			
	values	GAN	WGAN-GP	IWGAN-GP	
2016	53012.2906	49929.9152	49232.3567	51440.8984	
2017	53340.04034	49794.9574	49835.3661	51967.0522	
2018	54348.09278	50640.7640	50236.8975	52425.5705	
2019	54801.2475	50760.6806	50484.1072	52812.3368	
2020	56566.0091	50499.2780	50552.3718	53296.7498	
		TABLE V			
		YEARS 2016-2020 (THE THOUSAND TONS)		RELATIVE ERROR OF THE PROPOSED IWGAN-GP FOR THE	
Year			Relative Errors		
	GAN		WGAN-GP	IWGAN-GP	
2016	5.81%		7.13%	2.96%	

	Year	Relative Errors					
		GAN	WGAN-GP	IWGAN-GP			
	2016	5.81%	7.13%	2.96%			
	2017	6.65%	6.57%	2.57%			
	2018	6.82%	7.56%	3.54%			
	2019	7.37%	7.88%	3.63%			
	2020	10.73%	10.63%	5.78%			
			REFERENCES				
[1].			Computation, vol. 9, no. 8, pp1735-178, 1997.	S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory," Neural			
$\lceil 2 \rceil$.		preprint arxiv:1406.1078, 2014.		K. Cho, B. Van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," arxiv			
$\lceil 3 \rceil$.		H. Sen, "Time Series Prediction based on Improved Deep Learning," IAENG International Journal of Computer Science, vol. 49, no.4, pp1133-1138, 2022.					
[4].		J. Chen, H. Dai, S. Wang, and C. Liu, "Improving Accuracy and					

REFERENCES

-
- 2019 7.37% 7.88% 3.63%

2020 10.73% 10.63% 5.78%
 REFERENCES

[1]. S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory," Neural

Computation, vol. 9, no. 8, pp1735-178, 1997.

[2]. K. Cho, B. Van Merrienboer, C. Gulce 10.63% 5.78%

ID.63% 5.78%

ID.62% 7 Cemputer Memory, "learning, F. Bougares, H.C.

ID.50, The M 020 10.73% 10.63%

REFERENCES

S. Hochreiter, J. Schmidhuber, "Long Short-Term M

Computation, vol. 9, no. 8, pp1735-178, 1997.

K. Cho, B. Van Merrienboer, C. Gulcehre, D. Bahdana

H. Schwenk and Y. Bengio, "Learning phra **EXECT REFORE SEALL SET ALL SCELL SET ALL SCELL SET ALL COMPUTE COMPUTE SET ALL SCELUSE IN A SCHOTT REFORMATION**

1997. I. Chen, B. Van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk and Y. Bengio, "Learnin REFERENCES

S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory," Neural

Computation, vol. 9, no. 8, pp1735-178, 1997.

K. Cho, B. Van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares,

H. Schwenk and Y. Bengio, "Lea REFERENCES

S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory," Neural

Computation, vol. 9, no. 8, pp1735-178, 1997.

K. Cho, B. Van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares,

H. Schwenk and Y. Bengio, "Lea [1]. S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory," Neural

Computation, vol. 9, no. 8, pp1735-178, 1997.

[2]. K. Cho, B. Van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares,

H. Schwenk and Y. Bengio, "Lear S. Hochreiter, J. Schmidhuber, "Long Short-Term Memory," Neural
Computation, vol. 9, no. 8, pp1735-178, 1997.
K. Cho, B. Van Merrienboer, C. Gulcherle, D. Bahdanau, F. Bougares,
H. Schwenk and Y. Bengio, "Learning phrase r Computation, vol. 9, no. 8, pp1735-178, 1997.

K. Cho, B. Van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine [2]. K. Cho, B. Van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares,

H. Schwenk and Y. Bengio, "Learning phrase representations using

RNN encoder-decoder for statistical machine translation," arxiv

preprint arxiv:140
- H. Schwenk and Y. Bengto, "Learning phrase representations using

RNN encoder-decoder for statistical machine translation," arxiv

preprint arxiv:1406.1078, 2014.

H. Sen, "Time Series Prediction based on Improved Deep Lea RNN encoder-decoder for statistical machine translation," arxiv
preprint arxiv:1406.1078, 2014.
H. Sen, "Time Series Prediction based on Improved Deep Learning,"
H. Sen, "Time Stries Prediction based on Improved Deep Learn
-
-
-
- **IAENG International Journal of Comping 17**.

I. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,

S. Ozair, A. Courville and Y. Bengio, "Generative adversarial

networks," in Communications of the ACM, vo **IAENG International Journal of Compute**

I. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,

S. Ozair, A. Courville and Y. Bengio, "Generative adversarial

networks," in Communications of the ACM, vol. 6 **IAENG International Journal of Computer Sci**
I. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,
S. Ozair, A. Courville and Y. Bengio, "Generative adversarial
networks," in Communications of the ACM, vol. **IAENG Int**

I. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. 2

S. Ozair, A. Courville and Y. Bengio, "Genetworks," in Communications of the ACM,

139-144, 2020.

H. Lin, C. Chen, G. Huang and A. Jafari, "Stock p

generati **IAENG International Journal of Computer S**

[7]. I. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,

S. Ozair, A. Courville and Y. Bengio, "Generative adversarial

networks," in Communications of the ACM **IAENG International Journal of Computer Sci**

I. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,

S. Ozair, A. Courville and Y. Bengio, "Generative adversarial

networks," in Communications of the ACM, v **IAENG Internatio**

I. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Ward

S. Ozair, A. Courville and Y. Bengio, "Generative ac

networks," in Communications of the ACM, vol. 63, n

139-144, 2020.

H. Lin, C. Chen, G **IAENG International Journal of Compute**

[7]. I. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,

S. Ozair, A. Courville and Y. Bengio, "Generative adversarial

networks," in Communications of the ACM, **IAENG International Journal of Compute**

I. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,

S. Ozair, A. Courville and Y. Bengio, "Generative adversarial

networks," in Communications of the ACM, vol. 6 **EXECTME INCETRO INTERTOR INTERTOR INTERTOR INTERTOR IN COMPUTE

1. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,

8. Ozair, A. Courville and Y. Bengio, "Generative adversarial

139-144, 2020.

H. Lin,** [7]. I. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,

S. Ozair, A. Courville and Y. Bengio, "Generative adversarial

networks," in Communications of the ACM, vol. 63, no.11, pp

19. 144, 2020.

[8]. H. I. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,
S. Ozair, A. Courville and Y. Bengio, "Generative adversarial
networks," in Communications of the ACM, vol. 63, no.11, pp
139-144, 2020.
H. Lin, C. Chen,
-
- 2020.

[10]. S. B. Jiang, C. F. Bai, and C. F. Du, "Traffic forecast based on 1. Goodfellow I, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley,

S. Ozair, A. Courville and Y. Bengio, "Generative adversarial

networks," in Communications of the ACM, vol. 63, no.11, pp

139-144, 2020.

H. Lin, C. C S. Ozair, A. Courville and Y. Bengio, "Generative adversarial
networks," in Communications of the ACM, vol. 63, no.11, pp
139-144, 2020.
[8]. H. Lin, C. Chen, G. Huang and A. Jafari, "Stock price prediction using
generativ networks," in Communications of the ACM, vol. 63, no.11, pp

13-144, 2020.

H. Lin, C. Chen, G. Huang and A. Jafari, "Stock price prediction using

generative adversarial networks," Journal of Computer Science, vol

17, no 139-144, 2020.

H. Lin, C. Chen, G. Huang and A. Jafari, "Stock price prediction using

generative adversarial networks," Journal of Computer Science, vol

17, no.3, pp188-196, 2021.

J. Wang, H. Zou, D. Qu, and L. Bai. "F [8]. H. L.m, C. Chen, G. Huang and A. Jafari, "Stock price prediction using

enerative adversarial retworks," Journal of Computer Science, vol

17, no.3, pp188-196, 2021.

[9]. J. Wang, H. Zou, D. Qu, and L. Bai. "Financia generative adversarial networks," Journal of Computer Science, vol
17, no.3, pp188-196, 2021.
11, wang, H. Zou, D. Qu, and L. Bai. "Financial time series prediction
based on empirical mode decomposition generative adversar 17, no.3, pp1888-196, 2021.

1. Wang, H. Zou, D. Qu, and L. Bai. "Financial time series prediction

based on empirical mode decomposition generative adversarial

networks," Computer Applications and Software, vol. 37, pp29 [9]. J. Wang, H. Zou, D. Qu, and L. Bai. "Financial time series prediction
based on empirical mode decomposition generative adversarial
networks," Computer Applications and Software, vol. 37, pp293-297,
2020.
[10]. S. B.
-
- based on empirical mode decomposition generative adversarial

2020.

2020.

S. B. Jiang, C. F. Bai, and C. F. Du, "Traffic forecast based on

empirical mode decomposition and RBF neural network," Advanced

Materials Resear
-
-
- networks," Computer Applications and Software, v
2020.

S. B. Jiang, C. F. Bai, and C. F. Du, "Traffic

empirical mode decomposition and RBF neural ne

Materials Research, vol. 846, pp1270-1273, 2014.

I. Gulrajani, F. Ahm 2020.

[10]. S. B. Jiang, C. F. Bai, and C. F. Du, "Traffic forecast based on

empirical mode decomposition and RBF neural network," Advanced

Materials Research, vol. 846, pp1270-1273, 2014.

[11]. I. Gulrajani, F. Ahmed, S. B. Jiang, C. F. Bai, and C. F. Du, "Traftic forecast based on empirical mode decomposition and RBF neural network," Advanced Materials Research, vol. 846, pp1270-1273, 2014.

I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumo empirical mode decomposition and RBF neural network," Advanced

I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin and A. C.

Courville, "Improved training of Wasserstein gans," Advances in

neural information processing sys Materials Research, vol. 846, pp1270-1273, 2014.

I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumc

Courville, "Improved training of Wasserstein gan

neural information processing systems, vol. 30, pp1-

S. Takahshi, Y. Chen, [11]. I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin and A. C.

Couville, "Improved training of Wasserstein gans," Advances in

neural information processing systems, vol. 30, pp1-11, 2017.

[12]. S. Takahashi, Y. Chen, Courville, "Improved training of Wasserstein gans," Advances in
encural information processing systems, vol. 30, pp1-11, 2017.

S. Takahashi, Y. Chen, and K. Tanaka-Ishii, "Modeling financial

time-series with generative a neural information processing systems, vol. 30, pp1-11, 2017.

S. Takahashi, Y. Chen, and K. Tanaka-Ishii, "Modeling financial

time-series with generative adversarial networks," Physica A:

Statistical Mechanics and its A S. Takahashi, Y. Chen, and K. Tanaka-Ishii, "Mo
time-series with generative adversarial networks
Statistical Mechanics and its Applications, vol.527, 1:
B. Gong, and F. Y. Yin, "Measurement of the impact
on China's grain c tme-series with generative adversarial networks," Physica A:

Statistical Mechanics and its Applications, vol.527, 121261, 2019.

[13]. B. Giong, and F. Y. Yin, "Measurement of the impact of urbanization

on China's grain Statistical Mechanics and its Applications, vol.527, 121261, 2019.

20. Gong, and F. Y. Yin, "Measurement of the impact of urbanization

on China's grain consumption," Theory Pract. Finance for twhanziation

c. J. Ogbonna, B. Gong, and F. Y. Ym, "Measurement of the impact of troson China's grain consumption," Theory Pract. Financ. Econ. pp134-140, 2018.
C. J. Ogbonna, C. J. Nweke, E. C. Nwogu, and I. S. Iwueze, Transform as an Alternative to on China's grain consumption," Theory Pract. Financ. Econ, vol. 39,
pp134-140, 2018.
[14]. C. J. Qgbonna, C. J. Nweke, E. C. Nwogu, and I. S. Iwueze, "Wavelet
Transform as an Alternative to Power Transformation in Time Ser pp134-140, 2018.

C. J. Ogbonna, C. J. Nweke, E. C. Nwogu, and I. S. Iwueze, "Wavelet

C. J. Ogbonna, C. J. Nweke, E. C. Nwogu, and I. S. Iwueze, "Wavelet

Transform as an Alternative to Power Transformation in Time Series C. J. Ogbonna, C. J. Nweke, E. C. Nwogu, and I. S. Iwueze, "Wavelet
Transform as an Alternative to Power Transformation in Time Series
Analysis, "Bulletin of Mathematical Sciences and Applications, vol.
17, pp57-74, 2016.

- Transform as an Alternative to Power Transformation in Time Series
Analysis," Bulletin of Mathematical Sciences and Applications, vol.

[15]. W. Wang, K. Chau, D. Xu, and X. Y. Chen, "Improving forecasting

accuracy of ann Analysis," Bulletin of Mathematical Sciences and Applications, vol.
17, pp57-74, 2016.
W. Wang, K. Chau, D. Xu, and X. Y. Chen, "Improving forecasting
accuracy of annual runoff time series using ARIMA based on EEMD
decompo 17, pp57-74, 2016.

W. Wang, K. Chau, D. Xu, and X. Y. Chen, "Improving forecasting

accuracy of annual runoff time series using ARIMA based on EEMD

decomposition," Water Resources Management, vol. 29,

pp2655-2675, 2015.
-
-
-
- [15]. W. Wang, K. Chau, D. Xu, and X. Y. Chen, "Improving forecasting

accomposition," Water exists and SRIMA based on EEMD

decomposition," Water Resources Management, vol. 29,

pp2655-2675, 2015.

[16]. D. Song, H. Chen, accuracy of annual runoft time series using ARIMA based on EEMID
decomposition," Water Resources Management, vol. 29,
pp2655-2675, 2015.
D. Song, H. Chen, G. Jiang, and Y. Qin, "Dual stage attention based
recurrent neural decomposition," Water Resources Management, vol. 29,
pp.2655-2675, 2015.
D. Song, H. Chen, G. Jiang, and Y. Qin, "Dual stage attention based
recurrent neural network for time series prediction," U.S. Patent No.
10, 929, 67 pp2655-2675, 2015.

D. Song, H. Chen, G. Jiang, and Y. Qin, "Dual stage attention based

recurrent neural network for time series prediction," U.S. Patent No.

10, 929, 674. 23 Feb. 2021.

Y. Li, Z. Zhu, D. Kong, H. Han, a
- [16]. D. Song, H. Chen, G. Jiang, and Y. Qm, "Dual stage attention based
recenter neural network for time series prediction," U.S. Patent No.
10, 929, 674. 23 Feb. 2021.
[17]. Y. Li, Z. Zhu, D. Kong, H. Han, and Y. Zhao, " recurrent neural network for trime series prediction," U.S. Patent No.
10, 929, 674. 23 Feb. 2021.
Y. Li, Z. Zhu, D. Kong, H. Han, and Y. Zhao, "EA-LSTM:
Evolutionary attention-based LSTM for time series prediction,"
Knowl 10, 929, 674. 23 Feb. 2021.

Y. Li, Z. Zhu, D. Kong, H. Han, and Y. Zhao, "EA-LSTM:

Evolutionary attention-based LSTM for time series prediction,"

Knowledge-Based Systems, vol. 181, pp 1-9, 2019.

ISTM and BiLSTM in forc Y. L.1, Z. Zhu, D. Kong, H. Han, and Y. Z
Evolutionary attention-based LSTM for time s
Knowledge-Based Systems, vol. 181, pp 1-9, 2019
S. Siami-Namini, N. Tavakoli, and A. S. Namin, "T
LSTM and BiLSTM in forceasting time s
- Evolutionary attention-based LSTM for time series prediction,"

Knowledge-Based Systems, vol. 181, pp 1-9, 2019.

[18]. S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The performance of

LSTM and BiLSTM in forecasting tim Knowledge-Based Systems, vol. 181, pp 1-9, 2019.

S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The performance of

LSTM and BiLSTM in forecasting time series," 2019 IEEE

International Conference on Big Data (Big Data), S. Siami-Namini, N. Tavakoli, and A. S. Namin, "The performance of

LSTM and BiLSTM in forceasting time series," 2019 IEEE

International Conference on Big Data (Big Data), pp3285-3292, 2019.

O. S. Bérot, H. Canot, P. Dur LSTM and BiLSTM in forecasting time series," 2019 IEEE
International Conference on Big Data (Big Data), pp3285-3292, 2019.
O. S. Bérot, H. Canot, P. Durand, B. Hassoune-Rhabbour, H.
Acheritobehere, C. Laforet, and V. Nassi 2020.

[22]. K. F. Wang, C. Gou, Y. J. Duan, Y. L. Lin, and X. H. Zheng, [9]. O. S. Bérot, H. Canot, P. Durand, B. Hassoune-Rhabbour, H.

Achritobehere, C. Laforet, and V. Nassiet, "Choice of Parametrs of

an LSTM Network, based on a Small Experimental Dataset,"

Engineering Letters, vol. 32, n Acheritobehere, C. Laforet, and V. Nassiet, "Choice of Parameters of

an LSTM Network, based on a Small Experimental Dataset,"

Engineering Letters, vol. 32, no. 1, pp59-71, 2024.

D. Hussain, T. Hussain, A. A. Khan, S. A. an LSTM Network, based on a Small Experimental Dataset,"
Engineering Letters, vol. 32, no. 1, pp59-71, 2024.
D. Hussain, T. Hussain, A. A. Khan, S. A. A. Naqvi and A. Jamil, "A
deep learning approach for hydrological time-Engineering Letters, vol. 32, no. 1, pp59-71, 2024.

[20]. D. Hussain, T. Hussain, A. A. Khan, S. A. A. Naqvi and A. Jamil, "A deep learning approach for hydrological time-series prediction: A case study of Gilgit river ba D. Hussain, T. Hussain, A. A. Khan, S. A. A. Naqvi and A. Jamil, "A
deep learning approach for hydrological time-series prediction: A case
study of Gilgit river basin," Earth Science Informatics, vol.13,
pp^{915-927, 2020.} deep learning approach for hydrological time-series prediction: A case
study of Gilgit river basin," Earth Science Informatics, vol.13,
pp915-927, 2020.
F. Liu, X. Zhou, J. Cao, Z. Wang, T. Wang, H. Wang, and Y. Zhang."
An pp915-927, 2020.

[21]. F. Liu, X. Zhony, J. Cao, Z. Wang, T. Wang, H. Wang, and Y. Zhang."

Anomaly detection in quasi-periodic time series based on automatic

data segmentation and attentional LSTM-CNN," IEEE Transaction [21]. F. Liu, X. Zhou, J. Cao, Z. Wang, T. Wang, H. Wang, and Y. Zhang."

Anomaly detection in quasi-periodic time series based on automatic

data segmentation and attentional LSTM-CNN," IEEE Transactions

on Knowledge and
-
- data segmentation and attentional LSTM-CNN," IEEE Transactions

on Knowledge and Data Engineering, vol. 34, no. 6, pp2626-2640,

2020.

Chenerative adversarial networks: The state of the art and beyond, "

Generative adver on Knowledge and Data Engineering, vol. 34, no. 6, pp2626-2640,
2020. K. F. Wang, C. Gou, Y. J. Duan, Y. L. Lin, and X. H. Zheng,
"Generative adversarial networks: The state of the art and beyond,"
Acta Automatica Sinica, 2020.

[22]. K. F. Wang, C. Gou, Y. J. Duan, Y. (22). K. F. Wang, C. Gou, Y. J. Duan, Y. The Acta Automatica Sinica, vol. 43, no. 3, [23]. X. Zhou, Z. Pan, G. Hu, S. Tang, ar Prediction on High-Frequency Data U. Nets," Mat