UNIVERSITY OF LIVERPOOL

Microwave Based Multiphase Flow Detection Using Convolutional Neural Network

by

Jingyuan Jiang

A thesis submitted in accordance with the requirements for the award of the degree of *Master of Philosophy* of the University of Liverpool

September 2019

DECCLARATION

I hereby declare that the content of this article is original unless the work of others is specifically mentioned. The content of the paper is an original work that has never been published in any institution or obtained any other degree.

The copyright of this thesis rests with the author. Copies (by any means) either in full, or of extracts, may not be made without prior written consent from the author. Copyright © 2019 Jingyuan Jiang. All rights reserved.

Signed,

Jingyuan Jiang

Dedicated to my family and friends

ACKNOWLEDGEMENTS

ABSTRACT

It is of great interest and importance to detect the amount of water, oil and gas passing through a pipe in real time for many industrial applications. However, due to the different physical properties of the three fluids (water, oil, and air/gas), it is beyond most companies' ability to effectively and accurately calculate the fluid content. Other situations could also make such estimation more difficult. For example, at the industrial level, the internal environment of the pipe is often accompanied by high pressure while the liquid flows at high speed. Also, highly polluted oil can easily damage electronic equipment. Consequently, there has been constant concerns regarding to multiphase flow detection.

The existing advanced three-phase flow detection technology generally develops around imaging technology, such as electrical resistance tomography, electrical capacitance tomography and microwave tomography. The advantage of these methods is that the internal environment of the pipe can be reconstructed intuitively, thereby achieving the purpose of "perspective". However, the shortcomings of these methods are also obvious: firstly, the reconstruction operation is complicated, which requires considerable amount of calculation, potentially causing severe postponement. Secondly, when there is large amount of high-conductivity liquid (e.g., water) in the pipe, the amount of each fluid in the three-phase flow is hard to calculate.

In this thesis, a detection method combining microwave measurement and convolutional neural networks was proposed and validated. The author designed a microwave measurement multiphase flow experiment as well as a convolutional neural network model with microwave measurement data as input data. The Error Sum of Square of oil predicted content is 0.81. Since a trained model is applied to predict water and oil content, this system involves less calculation and can be considered almost real-time detection. Besides, this system is not only based on non-destructive testing, but also suitable for a wide range of materials, included the liquid with high water content inside the pipe.

LIST OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Compared to the dramatic development of innovation projects in industrial production, technology to accurately detect the effect of process tolerances on raw materials quality has already fallen behind. It takes several production steps to fully qualify a new product, which could be inconvenient and inaccurate. Unilever, one of the largest consumer goods producers in the world, addressed liquid quality detection as one of the three challenges at the beginning of 2018, which specifically refers to "Development of robust in-line measurement techniques to track microstructure evolution during processing of personal care liquids (shampoo, body wash, conditioners)" [1]. The old method to detect the quality of the finished product, normally, is just to test it straightly at the end of producing. However, this method has high risk in terms of cost. For example, the product cannot be sold because it does not satisfy the quality standards though it has finished all the production steps. Therefore, companies would like to find some real-time detecting methods to avoid unqualified products as much as possible to reduce wastes.

Fig. 1 The petroleum industry needs real-time monitoring of data from product mining

In petroleum industry, the methods of real-time product testing are necessary and desperately in need. It is common to see Three-phase flows of gas and two liquid phases in the production of oil. Usually, an oil and gas reservoir contain gas, oil and

water. Although oil and gas are wanted, underground water is not. When the oil is extracted from the reservoir, underground water is likely to flow into the well because it could penetrate the rock via pores easily. What is more, it is beneficial to inject water into the reservoir in order to increase recovery of the oil and gas when a well gets older [2]. Because wells that have been extracted for different length of time contain different yield of gas, oil, and water, detecting the percentage of each liquid in the output is of great importance in determining the life of oil wells and the production cost. Therefore, it is important to accurately measure the different liquid's content in multiphase flow during the production process to help companies successfully achieve maximum return on investment [3]. Oil and gas pipeline monitoring sensors have gradually become a key factor in determining the profitability of the oil industry.

Fig. 2 Oil and gas pipeline monitoring sensors have become a new way to increase profits in the oil industry.

There have been continuous discussion and proposals regarding to multiphase flow/liquid testing, due to the complex and constant changing flow conditions, as well as the great difference in the dielectric constant between water, oil and air (around 80, 2 and 1 respectively) (Section 2.1).

During this process, many solutions have been proposed and various kinds of theories have been put forward. Some are intrusive testing methods, such as Time Domain Reflectometry (Section 2.1.1) and Wire-mesh Sensor (Section 2.1.2). The

others are non-intrusive testing methods like electrical capacitance tomography method (Section 2.1.3.2) and microwave tomography method (Section 2.1.4). Also, Since the product is multiphase, it could be separated into single phases of liquid by a large separator and then measured by different flow meter [4]. This kind of method integrated many separate liquid sensors and combined these liquid sensors into a multiphase flowmeter.

The solutions shown above could effectively solve the problem of measuring the liquid content in a certain state in the pipeline, but a common limitation is also obvious. In the actual production process, the multiphase flow is continuously flowing inside a tube/pipe. It means that the content of the liquid changes in every moment, and these sensors need to calculate the result repeatedly. Since most of these methods are based on complex mathematical operations, which require time to calculate, making it difficult to achieve real-time detection. Thus, a method to achieve real-time detection by reducing the amount of calculation is necessary.

In modern daily life, machine learning technology has been widely used: from web search and network filtering, to analyze network users' preferences as well as speech and image recognition. Machine learning technology appears more frequently in e-commerce, Internet products and other fields. This technology analyzes and summarizes network information, also predicts the results according to the needs of the users as well as the product information that users may be interested in [5].

Nowadays, most of the existing Neural Network (NN) (Section 2.2) models are based on the highly simplified brain dynamics. These powerful computational models have been applied in solving the problems of complex pattern recognition, function estimation as well as classification issues [6]. Also, there is no need to understand or analyze NN with a good knowledge of physical phenomena [7]. Deep Learning (DL) is part of NN, which uses a mathematical model that consists of multiple processing layers to learn samples with high-dimensional data. It uses the backpropagation (BP) algorithm to adjust the parameters in order to discover the complex internal structure of big data. DL shines in processing images, video, speech and audio. One of the typical methods of deep learning is convolutional neural networks (CNN) (Section 2.2.2). It has been applied effectively in lots of

fields. For example, 'Alpha go', an intelligent chess system that is popular all over the world, which is also based on CNN.

Fig. 3 'Alpha go' battles and defeats Korean Go players

However, it is a challenge to combine deep learning with industrial measurement. Since machine learning is not a reliable measure in the minds of many people who hold traditional ideas, people tend to use computational methods to make quantitative measurements. Although some people are not willing to try this method, there is an unavoidable tendency for combining artificial intelligence with industrial development. This thesis is going to study the possibility of using convolutional neural networks instead of traditional calculation methods to measure and predict liquid content in multiphase flow inside a pipe.

CHAPTER 1 LITERATURE REVIEW

1.1. Multiphase Flow Testing

Multiphase flow/liquid is a common raw material or output in many industrial fields, such as energy, chemical engineering and food engineering. One of the most typical multiphase flow testing problems is gas-oil-water three-phase flow testing, which is a measurement problem that caught the attention of the oil industry.

Facing the multiphase flow testing, the main difficulties are as following:

a. Great difference of permittivity between oil, gas and water.

b. Changing liquid content is at any time.

c. Complex liquid structure.

d. High pressure and temperature in pipeline.

The future market of the three-phase flowmeters is important if the application of three-phase flowmeters has a real goal to make a flowmeter perpetually be fitted into each individual well in all new surface and subsea developments [8].

Several methods for measuring multiphase flow are described below. Some of them are not suitable for measuring multiphase liquid in pipe, and others have been researched and expanded in decades. But all of them are typical methods in the measurement of multiphase flow.

1.1.1. Time Domain Reflectometry

Time domain reflectometry (TDR) is a pulse sampling technique. It is a relatively common method in the detection of cable impedance variations and continuity [9]. For the convenience of measurement, the transmission cable can be regarded as a series of circuit components, so the coaxial cable used for TDR measurement can be considered as an ideal transmission circuit, which is illustrated as Fig. 1.1.
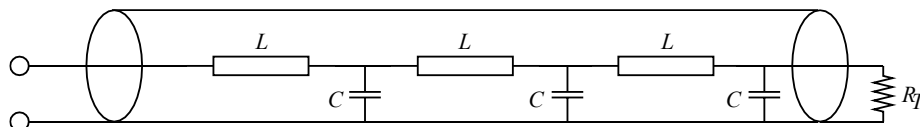
Fig. 1.1 Discrete model of an unbalanced, lossless transmission line.

The voltage-current relationship of a time-varying signal inside the length of the coaxial cable is decided by characteristic impedance, pulse propagation velocity, as

well as the reflection coefficient [10]. The characteristic impedance Z_0 depends on the capacitance C and inductance L per unit length:

$$Z_0 = \sqrt{\frac{L}{C}}$$

Z_0 is measured in ohms, L in henrys and C in farads. For an ideal model, the impedance can be considered as a pure resistor. The propagation speed v of the pulse signal depends on the capacitance and inductance per unit length:

$$v = \frac{1}{\sqrt{LC}}$$

All changes of the transmission line, such as changing material or shape, will result in the changes in impedance. The electromagnetic signal propagating on the transmission line will also partially or even totally reflect at the discontinuity of the impedance. If there is a coaxial cable terminated with a load Z_T, the reflection coefficient ρ could be calculated as:

$$\rho = \frac{Z_T - Z_0}{Z_T + Z_0}$$

The simplest TDR model is consisted by a pulse generator and pulse sampler which is shown in Fig. 1.2.
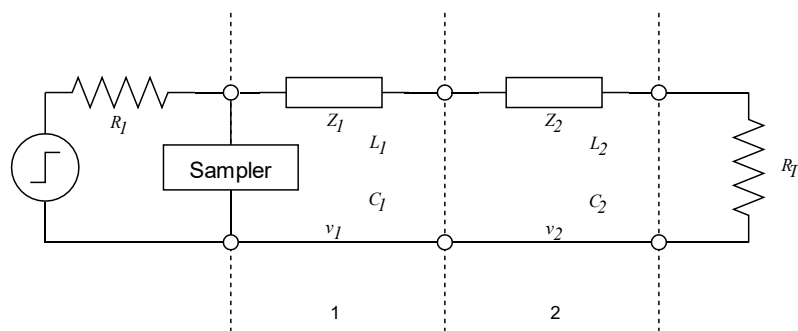
Fig. 1.2 TDR and two segment coaxial networks

This model shows a TDR, a transmission line which is made up by two series coaxial cables as well as a load resistor R_T. Two coaxial cable segments have its own impedance Z_1 and Z_2 ($Z_1 \neq Z_2$) which is given by their unit length capacitance (C_1 and C_2) and inductance (L_1 and L_2).

6
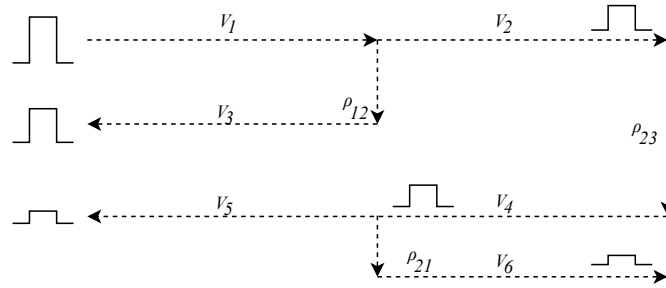
Fig. 1.3 Pulse reflection diagram

TDR could launch a pulse electrical signal. This pulse starts to ambulate on segment 1 with voltage amplitude V_1 and propagation velocity v_1 on t_0, at the time it arrives the linking with coaxial cable segment 2, then another pulse with amplitude V_2 will keep transmitting coaxial cable segment 2, however, with a new velocity. Meanwhile, a pulse with amplitude V_3 reflects along segment 1 to the sampling circuit. Reflection coefficient, ρ_{12}, is applied to calculate the amplitude produced by the transmitted and reflected pulses. At t_1, the reflected pulse will reach out the sampling circuit. Pulse V_3 keeps travelling along segment 2 and it would not stop until it arrives the termination resistor where it reflects. This reflected pulse has amplitude V_4, which is decided by the reflection coefficient named ρ_{23}. Then, this pulse transmits back along segment 2 until it arrives the connection with segment 1. Over again, a part of the pulse continues to travel to the sampling circuit that arriving at t_1 and the termination will be reflected another portion of pulse. V_5 and V_6 show these pulses. Finally, the amplitudes are decided by reflection coefficient called ρ_{21}. This process could be illustrated by Fig. 1.3.

This technology was later developed to measure the liquid level of different fluid. In 1960s, it was used to measure a single level and named "TDR liquid level sensor" [11]. In recent years, TDR technology is still receiving the attention of many people from all over the world. In 2012, Andrea Cataldo explored the

Fig. 1.4 MG1-E Single Level Probe for Liquids and Solids

use of a low-cost portable TDR device for the automatic control and real-time monitoring of the flow and the liquid level in intravenous (IV) medical infusions [12]. E. I. Trenkal (2017) proposed a solution based on TDR to determine the structure and layer parameters of multiphase liquids in real-time mode in order to improve the accuracy of measurement [13]. After decades of improvement, the level gauge based on TDR technology can now be mass-produced. Fig. 1.4 shows an example. The structure of TDR liquid level sensor could be described as Fig. 1.5.
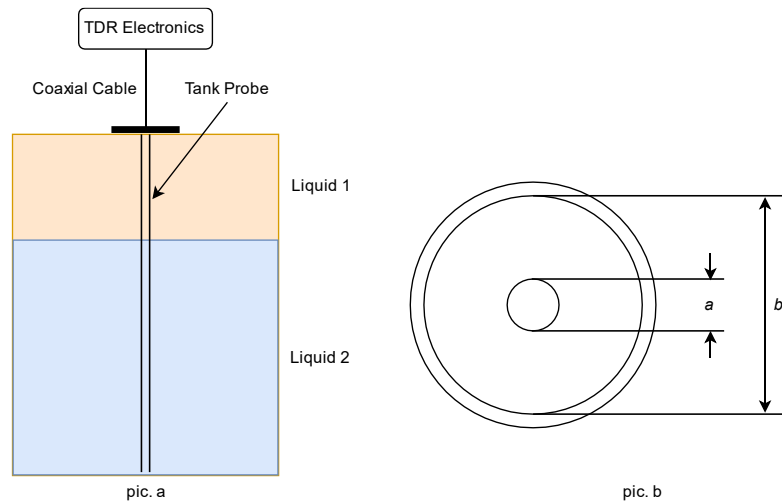
Fig. 1.5 The structure of TDR liquid level sensor (pic. a) and probe (pic. b)

The probe of TDR liquid level sensor is an open transmission line and the liquid inside the tank becomes the transmission line's insulation dielectric [9]. At the liquid interface, different dielectric constants of the fluid will cause changes in the impedance of the transmission line and some of the pulse signal to be reflected to the TDR receiver too. The liquid levels and the location of interface could be accurately determined by reflected pulse amplitude and time.

Compare with other fluid level sensors, TDR technology have a lot of advantages, such as simple mounting, high resolution, ability to measure multiple fluid interface levels and handle complex tank geometries. What is more, the dielectric constants of the fluid layers must improve with the depth along the investigation [9].

1.1.2. *Wire-mesh Sensor*

Wire-Mesh Sensor (WMS) is one of the imaging devices which can provide flow images at high spatial and temporal resolutions [14]. In many researches about WMS, it is mostly used in the decision of the distribution of the phase fraction and

visualization of the flow behavior within the cylindrical pipe. Flow structures that appear inside pipes with some flow rate fluids are usually analyzed [15]. Furthermore, take a typical situation as an example-when the fluid has two different liquids model or two-phase flow. This can be a blend of gas as well as oil that flow from a well when in the oil industry [16].
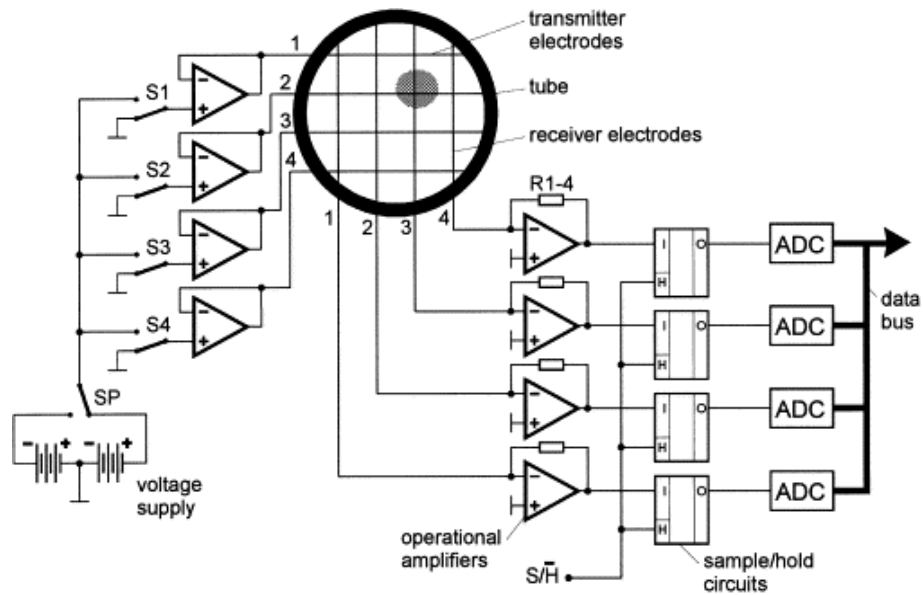
Fig. 1.6 Simplified scheme of the electrode-mesh device [17].

There are two models of working principle of the wire-mesh sensor, one is on the basic of the measurement of conductivity; the other is the permittivity value in the research area. In 1998, Prasser firstly designed a wire-mesh sensor which based on conductivity measurements and Fig. 1.6 shows the electrode-mesh sensor structure in his paper [17].
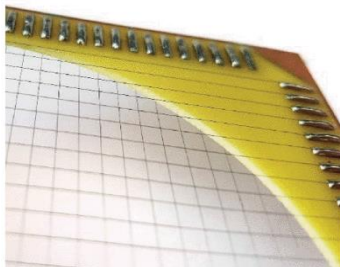
This WMS has two layers of conductive wire grids placed in parallel with a spacing of 2 mm and their wires are arranged to intersect at 90 degrees in space. Fig. 1.7 shows the layout of transmission lines in wire-mesh sensor [18]. When the spacing of the two planes is occupied by the conductive phase, the voltage pulse of the first layer of the wire grid (emission plane) reaches the second

Fig. 1.7 Layout of transmission lines in wire-mesh sensors [18].

9

plane (receiving plane). In 2015, Weiling Liu improved the conductivity wire-mesh sensor measurement system to test the horizontal gas-water two-phase flow and achieved 200 frames/s time resolution and 3.1 mm spatial resolution [19]. However, this kind of sensor is not suitable for all multiphase flow problems due to the need to have a conductive phase [20]. In 2007, on the basic of permittivity measurement, a wire-mesh sensor was exploited for visualizing and investigating of non-conducting fluids. Gas-oil two-phase flow is a specific example [21].

Compared to other tomography systems, the disadvantages of WMS as an intrusive detection method can be partially compensated by high temporal resolution, low cost, and accurate local measurements rather than parameter reconstruction [16]. Therefore, WMS is often used to compensate for the shortcomings of electrical tomography process [15].

1.1.3. Electrical Tomography

Tomographic imaging is a technique which helps to reconstruct the internal distribution of an object with a physical parameter. This technique used an image with gray or color scale for easy visualization to present the results. Tomographic imaging provides a special imaging method to the oil, chemical (Fig. 1.8), pharmaceutical as well as food industries to make the industrial processes to be visualized and provides great information for control and optimization without invading the pipes and vessels instead of a single-point measurement.

Fig. 1.8 Electrical resistance tomography for kids with Cystic Fibrosis in 2019.

On the basic of gaining information on electronic impedance which includes resistance, capacitance as well as inductance, electronic tomography is applied in electrical property. For example, it can be used in the description of conductivity, permittivity and permeability and distribution of materials [22]. In addition, the property of the material as well as a specific method are needed to measure the electrical property. Take the oil industry as an example, gas/oil/water multiphase is full of challenges. Conductivity and permittivity can be regarded as the dominant electrical property. Electrical capacitance tomography which is known as ECT has been researched and developed for a long time. It has been known that the ECT can provide quantitative information about the distribution of materials with an electrical dominance in the way of cross-sectional. As for the shortage of electrical capacitance tomography, the fluid inside the pipe must be continuous oil [23]. It means ECT cannot work well when the water is a continuous phase because of the media is conductive dominance. Therefore, in the case of water, the conductivity of water limits the capacitance electrodes so that the values capacitance couldn't be measured precisely [22]. On the contrary, electrical resistance tomography (ERT) could provide information about conductive distribution. For example, the inter-electrode conductance can be measured even if the water is a continuous phase. For a multiphase flow meter, it is more significant to measure oil-continuous flows and water-continuous flows with more water is produced with oil and gas.

1.1.3.1. *Electrical Capacitance Tomography*

In electrical capacitance tomography, the internal dielectric constant distribution structure is reconstructed by measuring the external capacitance between the electrode plates. These plates are mounted around the outer perimeter of the object. During the measurement, a voltage is applied to the electrodes according to a predetermined sequence. A single electrode is excited in sequence and the capacitance between all other electrode pairs is measured simultaneously [24]. Fig. 1.9 shows a typical 12-electrode ECT sensor. These electrodes are mounted on the outside of the insulating tube or container. Different electrode pairs can detect the cross-section of the liquid in the pipe or vessel from different angles. A complete measurement cycle has a total of $N(N-1)$ independent measurements, where N is the number of electrode plates. In the ECT sensor, 66 independent measurements

can be made in one cycle.
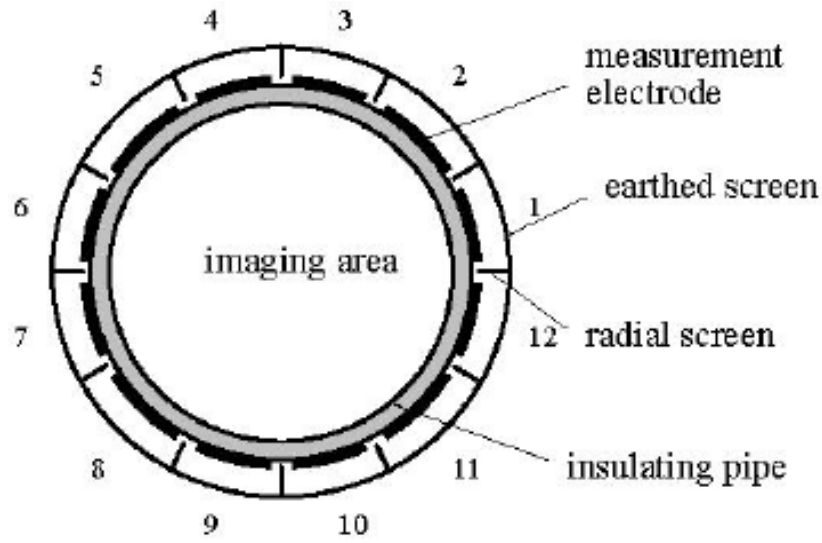
Fig. 1.9 Typical ECT sensor with 12 sensing electrodes [24]

There are two types of circuits for ECT sensors, one is a charge/discharge circuit [25] and the other is an AC-based circuit. Fig. 1.10 is the charging/discharging circuit. The output of this circuit is [26]

$$V_0 = 2fV_cC_xR_f + e_1 - e_2$$

Where e_1 and e_2 are the offset voltages of op-amps.
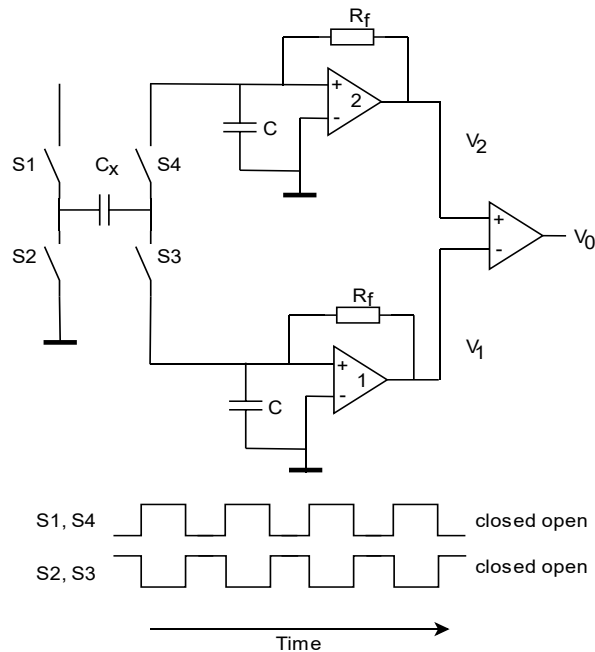
Fig. 1.10 Charge/discharge circuit in ECT system

The advantages of this circuit are linear, low cost and high sensitivity (0.56 V/pF)

12

[24]. But it has the disadvantage of charge injection of CMOS switches [27].

Base on sine-wave excitation and phase sensitive demodulation, an improved circuit called the AC circuit is designed in Fig. 1.11. This circuit uses DDS signal generators with programmable amplitude, frequency and phase. Therefore, it is available to test both capacitance and loss conductance with this circuit [24].
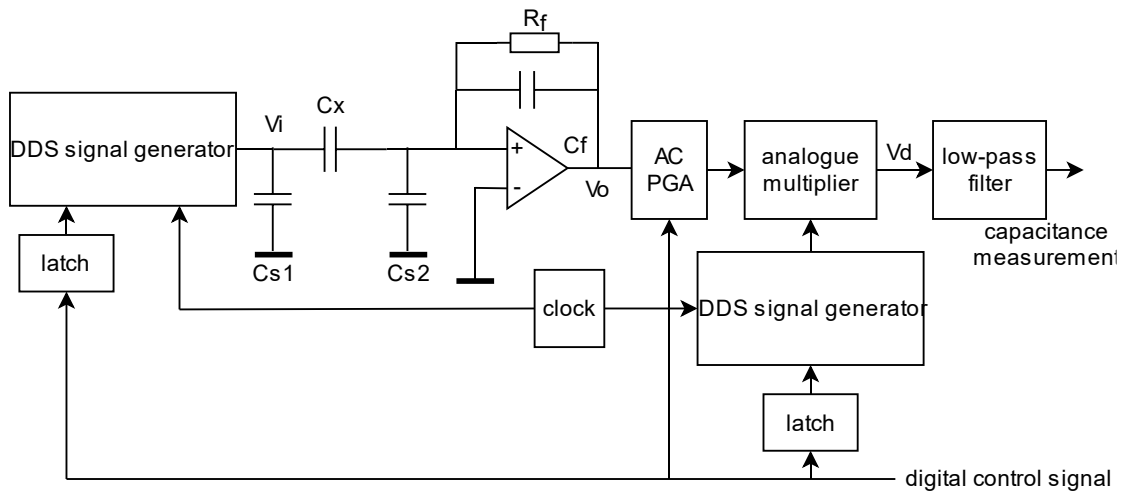
Fig. 1.11 AC-based circuit employed in ECT system

After all the electrode pairs have measured the capacitance, the measured capacitance values are reconstructed into an image as a dielectric constant distribution. In this process, it is necessary to normalize the capacitance measurements. In 2011, Zhiheng Guo compared the capacitance measurement normalization in different methods in the ECT system and studied the relationship between image quality and normalization method selection. He believed that optimizing the normalization method can improve the performance of some iterative algorithms [28].

ECT is merits-less costly, not radiate, quick-responsive, and non-instructive as well as non-invasive when compared with other modalities. What is more, ECT can be applied in strict environment with great pressure and high temperature. In the 1980's [29], the US Department of Energy in Morgantown had the very first try to use ECT in the process of visualizing and industrial process. Furthermore, Electrical Capacitance Tomography (ECT) is one of the important industrial applications that first developed and the most mature tomographic imaging technique.

The electrical resistance tomography (ERT) has been researched for decades and become one of the most significant process tomography [30]. It also has wide uses in many different industrial fields, such as metallurgy, petroleum, paper making as well as environmental engineering [31].
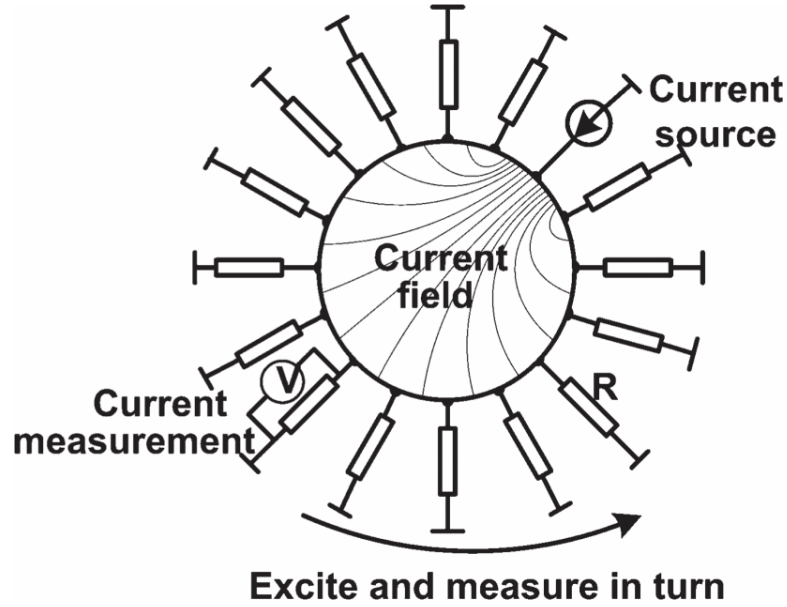
Fig. 1.12 Single drive electrode method [32].

Fig. 1.12 is an ERT schematic based on a single drive electrode method [32]. After a constant current is injected into one of the electrodes, the current is continuously measured from the other electrodes. The excitation current is then injected into the next electrode to repeat the above steps. A total of 16 sets of measurements are repeated for all independent data and each set containing 15 data points. Therefore, the number of outputs is 15×16=240. The resistor R in the figure is the grounding connection resistance, and the voltage of this resistor R can be measured. Based on the above data, the Ohm's law can be used to derive the output current.

The following Laplace equation expresses the relationship between potential distribution, conductivity and the applied current in this method.

$$\nabla \times (\sigma \nabla \varphi) = 0, \text{ between electrodes}$$

$$\int_{s} \sigma \times \frac{\partial \varphi}{\partial n} ds = +I, \text{ for injected electrode}$$

$$\int_{s-} \sigma \times \frac{\partial \varphi}{\partial n} ds = -I_x, \text{ for measured electrodes.}$$

In addition, the exciting current and the output current meet the following equation:

$$\sum_{i=1}^{k} I_i = I$$

where φ represents the electric potential, σ represents the conductivity distribution in the field, I represents the exciting current, I_i represents the output current, and n is the outer normal vector of each point at the boundary of the sensing field.

As the new generation of process parameter measuring technique, it can be known as an on-line measurement tool with the merits of visual, on-instructive, less costly as well as non-radiate. What is more, based on the fact that different conductivity with different medium varies, ERT proved the fact that when acquired conductivity distribution of the sensing filed; the medium distribution of the measured field can be identified. The data acquisition strategy like the adjacent method, the adaptive method, the opposite method as well as the multireference method can decide the resolution and the accuracy of the reconstructed image [32].

The leading-edge measurement method uses ECT and ERT in combination for imaging called ECT/ERT dual-modality system for imaging. This system uses two separate electrode planes, one performing an ECT measurement and the other performing an ERT [33]. However, this method is an expedient derived from the inability to simultaneously measure high permittivity liquid and low permittivity liquid. It has two drawbacks. One is the need for two separate switches and data acquisition systems. The other is that it takes time for liquid to flow from one plane to another, so two separate electrode planes produce inconsistent images [34].

Yang *et al.* proposed an ERTv/ECT system based on a dual modality impedance analyzer with 4-wire sensing [23]. Here, the ERTv denotes electrical resistance tomography with voltage excitation. This system integrates the ERT and ECT sensors on the same plane, enabling simultaneous measurement of capacitance and resistance. Fig. 1.13 shows the structure of this system [23].
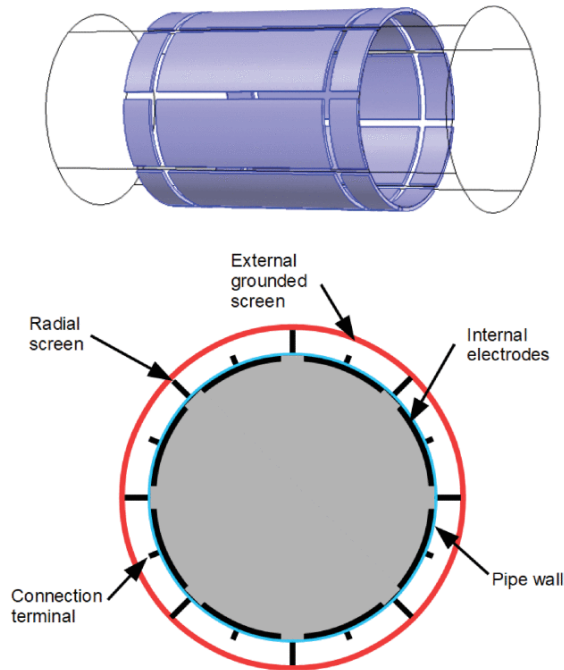
Fig. 1.13 3D sensor structure, 8 internal electrodes, axial grounded guards (top). Radial grounded guards, grounded screen and connection terminals in cross-sectional view (bottom) [23].

This ERTv/ECT system allows capacitance and resistance measurements to be obtained in a spatially and temporally coherent manner. It allows the measurement of multiphase flow over the entire range of water/liquid ratios. However, its imaging accuracy is not high enough.

1.1.4. *Microwave Tomography*

With comparison to the tomographic imaging techniques stated above, microwave tomography (MWT) generates images of dielectric properties on the basic of the evaluation of the scattered electromagnetic fields which is produced by an object [35]. When tracing the progress of the microwave tomography systems, it can be started in the late 1970s and early 1980s [36], [37]. In industrial process imaging, microwave tomography has different demands from that for medical imaging. Apart from spatial resolution, high temporal resolution as well as real-time imaging is significant for processes, flows or reactions that in high speed [35]. In a specific application, two different imaging methods – both quantitative and qualitative may be needed. The quantitative imaging methods are enough for many applications of representing distributions, patterns, or shapes. While the qualitative methods would

be more informative, it provides images that contrasted with quantitative methods that mentioned above and permittivity values from which other physical parameters. For example, qualitative imaging method derive data from density, moisture contend as well as phase fraction [35].
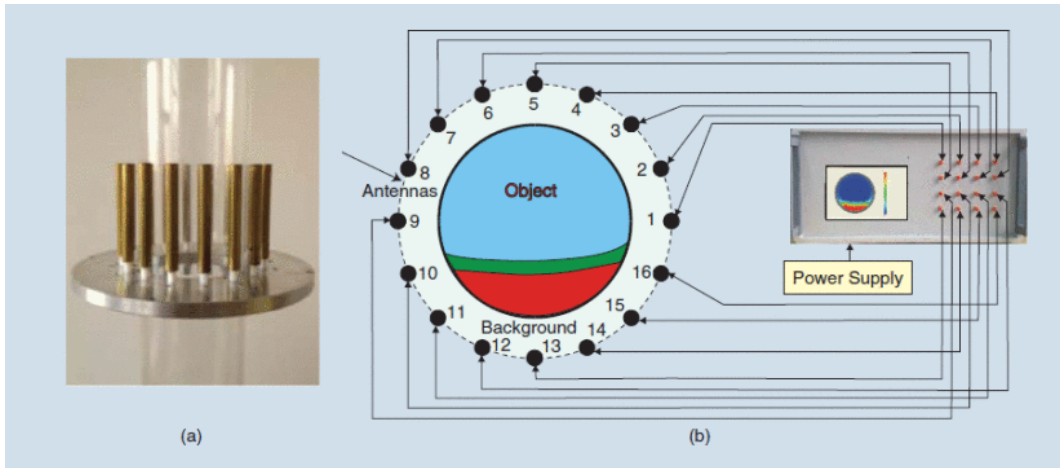
Fig. 1.14 The microwave tomography system setup [35]

The basic principle of microwave tomography is to measure the data of the scattered microwave field around the object with microwaves incident from different angels, and reconstruct the image as well as the multiple views of the object according to the dielectric properties (such as dielectric constant ε_r or dielectric contrast s). In a hardware for a microwave tomography system, circuits for microwave signal generation and detection, antennas for microwave signal transmitting and receiving as well as a personal computer are essentially needed. In this system, the antennas are usually placed in a background medium and this medium could be a matching medium of natural air. Fig. 1.14 shows the microwave tomography system setup [35].

When implementing microwave imaging, it is necessary to solve the forward-scattering and inverse-scattering problem. A two-dimensional (2D) model is a compromise solution to the problem of forward scatter. Microwave scattering is usually a three-dimensional scattering problem. It has been shown that the three-dimensional model can be reduced to a 2D model using a two-dimensional approximation [38], [39]. Numerous tomographic image reconstruction algorithms have been developed to solve the inverse-scattering problem. Such as iterative image reconstruction algorithms based on the Gauss-Newton method [40], the

Newton-Kantorovich method [41], the quasi-Newton method [42], the conjugate gradient method [43], and the sequential quadratic programming method [44]. In practical applications, these methods need to be selected according to the actual situation, because they all have different complexity, produce images of different quality, and different sensitivity to data.

1.1.5. Summary

The table below shows all the solutions mentioned above and summarizes their advantages and disadvantages.

Table 1.1 The advantages and disadvantages of the solutions mentioned above.

Solutions	Non-distractive	Imageable	Low error	Low computational complexity	Measure non-conductive multiphase flow	Measure highly conductive multiphase flow
TDR	×	×	√	√	√	√
WMS	×	√	√	√	√	√
ECT	√	√	√	×	√	×
ERT	×	√	√	×	×	√
ERTv/ECT	×	√	×	×	√	√
MWT	√	√	×	×	√	√

As can be seen from the above table, these six solutions for multiphase flow measurement cannot be characterized by non-destructive testing, entire range of water/oil ratios measurement, and low error. Therefore, it is necessary to research a method for measuring multiphase flow with these three characteristics.

1.2. Neural Network

Most machine learning can be divided into supervised learning and unsupervised learning. Supervised learning is a task of machine learning that learning a function that maps an input to an output on the basic of example input-output pairs [45]. On the contrary, unsupervised learning is a type of machine learning algorithm that applied in drawing inferences from the datasets that consist of data input without labeled responses. However, the early form of neural network architectures did not learn but only a few years later, the first ideas about unsupervised learning were developed. However, neural networks have been around even longer in a sense because the early supervised neural networks were originally variants of linear regression methods dating back at least to the early 1800s [46].

A standard neural network (NN) can consist of many simple, connected processors called neurons and each of them produces a sequence of real-valued activations. In this process, the input neurons can be activated by sensors perceiving the environment; other neurons are activated through weighted connections from the neurons activated previously. Depending on the issue and how the neurons are connected, some neurons may influence the environment by triggering actions, but it requires long causal chains of computational steps, where each step transforms (usually in a non-linear way) the collected activation of the network. Deep learning about assigning credit across many such steps accurately [46].

For ease of organization and understanding, only the basic backpropagation neural network and convolutional neural networks in deep learning are introduced here.

1.1.1. Backpropagation Neural Network

In the early period of pattern recognition, researchers attempted to replace hand-designed features with a trainable multi-layer network [47]. But this solution was not widely understood until 1980s. It has been recognized that a simple stochastic gradient descent can train a multi-layer architecture [5]. The backpropagation procedure can be used to calculate the gradient because the inputs and internal weights of the module are relatively smooth.

$$y_l = f(Z_l)$$
$$Z_l = \sum W_{kl}\, y_k$$
$$k\ \varepsilon\ H2$$

$$y_k = f(Z_k)$$
$$Z_k = \sum W_{jk}\, y_j$$
$$j\ \varepsilon\ H1$$

$$y_j = f(Z_j)$$
$$Z_j = \sum W_{ij}\, y_i$$
$$i\ \varepsilon\ Input$$
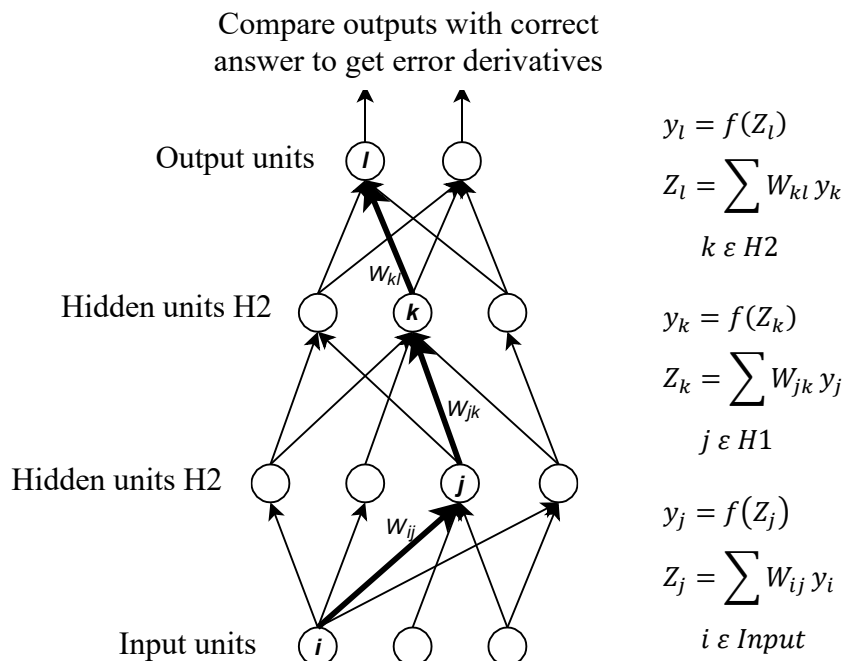
Fig. 1.15 The forward transfer in neural network and the equations

Fig. 1.15 shows the neural network forward transfer (two hidden layers and one output layer) and the equation for calculating the forward channel. Each layer of the neural network constitutes a module and the gradient can be propagated back through the module. The total output Z of each unit at each layer needs to be calculated first because it is the weighted sum of the cell outputs in the next layer. Then Z will be applied a nonlinear function to obtain the output of the unit. Bias terms are omitted for simplification.

In neural networks, there are a lot of nonlinear functions which are commonly used include rectified linear units (ReLU) (1.1), Sigmoids (1.2) and logistic functions logistic (1.3).

$$f(z) = max(0, z) \tag{1.1}$$

$$f(z) = 1/(1 + e^{-x}) \tag{1.2}$$

$$f(z) = \frac{1}{1 + exp\,(-z)} \tag{1.3}$$

Fig. 1.16 shows the backward transmission of the neural network. The output error derivative of each unit at each hidden layer is calculated, equal to the weighted sum of the error derivatives relative to the total input of the units in the upper layer. Then, the error derivative is converted from a derivative relative to the output into a derivative relative to the input. The method here is to multiply it by the gradient of $f(z)$.

$$\frac{\partial E}{\partial y_l} = y_l - t_l$$

$$\frac{\partial E}{\partial z_l} = \frac{\partial E}{\partial y_l} \frac{\partial y_l}{\partial z_l}$$

$$\frac{\partial E}{\partial y_k} = \sum_{l\,\varepsilon\,out} W_{kl} \frac{\partial E}{\partial z_l}$$

$$\frac{\partial E}{\partial z_k} = \frac{\partial E}{\partial y_k} \frac{\partial y_k}{\partial z_k}$$

$$\frac{\partial E}{\partial y_j} = \sum_{k\,\varepsilon\,H2} W_{jk} \frac{\partial E}{\partial z_k}$$

$$\frac{\partial E}{\partial z_j} = \frac{\partial E}{\partial y_j} \frac{\partial y_j}{\partial z_j}$$
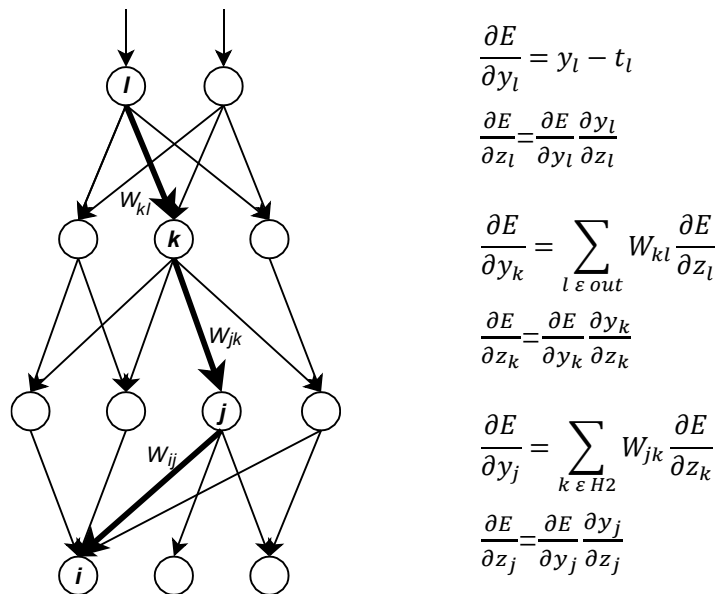
Fig. 1.16 The backward transfer in neural network and the equations

At the output layer, the cost function is differentiated to calculate the error derivative with respect to the output of a unit. If the cost function for unit l is $0.5(y_l - t_l)$, where t_l is the target value [48].

Feedforward neural network architectures, which learn to map a fixed-size input such as an image to a fixed-size output like a probability for each of several categories has been used for many applications of deep learning. A set of units calculates a weighted sum of their inputs from the layer before and pass the result by a non-linear function for the aim to go from one layer to the next.

Neural nets and backpropagation were abandoned by the community of machine-learning and ignored by the computer vision and speed-recognition during the late 1990s. It was a world-wide recognition that learning effective, multistage and feature extractors with little prior knowledge was infeasible. However, there was a special type of neural network which was deep as well as feedforward. Compared with networks that fully connected between adjacent layers, this kind of network was more easy-training and better-generalized. The network stated above was the convolutional neural network (ConvNet) [49], [50]. During the time when neural networks were disliked by people, ConvNet achieved many practical successes and recently it has been largely accepted by the computer-vision community.

1.1.2. *Convolutional Neural Network*

ConvNets are specific in the process of data that come in the way of multiple arrays. For example, a color image consists of three 2D arrays including pixel intensities in three color channels. Actually, many data modalities are in this kind of form (multiple arrays) such as 1D for signals and sequences, including language; 2D for images or audio spectrograms as well as 3D for video or volumetric images. Furthermore, local connections, shared weights, pooling and the use of many layers are the four key ideas behind ConvNets that benefit from the properties of natural signals.

A series of phases form the classic architecture of ConvNet, in which the convolutional layer and the pooling layer formed the first few stages. The convolutional layer takes the important role to detect local conjunctions of features from the previous layer. A local patch in the feature map of the previous layer

connected to each unit in the convolutional layer. A set of weights called filter banks are used to connect these units and patches. Then, functions such as ReLU make these locally weighted sums non-linear. Discrete convolution is the mathematical term for the filtering operation performed by the feature map. The name of the convolutional layer is also derived from this.

The pooling layer takes the responsibility to mix semantically same features into one. The reliable detection of the motif can be done by coarse-graining the position of each feature for the reason that the relative positions of the features forming a motif can vary somewhat. A typical pooling unit calculates the maximum of a local patch of units in one feature map or in some feature maps. Followed by more convolutional and fully connected layers, two or three steps of convolution, non-linearity and pooling are stacked. Backpropagating gradients by a ConvNet is as simple as by a regular deep network which allows all the weights in all the filter banks to be trained. For a detailed introduction to the convolutional neural network structure, please refer to Section 2.2.

The classic notions of simple cells and complex cells in visual neuroscience [51], and the overall architecture is reminiscent of the LGN–V1–V2–V4–IT hierarchy in the visual cortex ventral pathway are the direct inspiration of the convolutional and pooling layers in ConvNets [52]. Neocognitron, which is rooted by ConvNets, to some extent, has similar architecture with ConvNets but did not have an end-to-end supervised-learning algorithm such as backpropagation [53]. The recognition of phonemes and simple words was done by a primitive 1D ConvNet named a time-delay neural net [54].

When dating back to the early 1990s, there have been many applications of convolutional networks and it started with time-delay neural networks for the recognition of speech [55] and the reading of document [50]. The application of a ConvNet trained jointly with a probabilistic model that implemented language constraints was used in the system of document reading. 10% of all the cheques in the United States were read by this reading document system by the late 1900s. Later, Microsoft developed a number of ConvNet-based systems to recognize the optical character and handwriting [56]. What is more, in the early 1990s, ConvNets

were also tried for the detection of object in natural images which include faces as well as hands [57], it is also applied in the recognition of face [58].

Similar to the scenario that CNN has applied, multiphase flow detection is a task that requires a lot of complex calculations. According to Section 1.1.5, the proposed work should have the following characteristics: non-invasive, low error, low computational complexity, and suitable for entire range of water/oil ratios. Therefore, the possibility of microwave based multiphase flow detection using convolutional neural network is considered.

CHAPTER 2 METHODOLOGY

As presented in the literature review, it is difficult to calculate the liquid content of the three-phase flow by the conventional calculation method. The current mainstream three-phase flow measurement method is tomography. However, it is impossible to accurately measure the content of water and oil at the same time, especially in liquid environments with a lot of water as there is prominent difference in water and oil permittivity. The leading-edge measurement method of three-phase flow uses ECT and ERT in combination for imaging, which is called ECT/ERT dual-modality system for imaging. However, this method is an expedient due to the inability to simultaneously measure high permittivity liquid and low permittivity liquid.

Different from the above methods, the theoretical basis of measurement in this thesis is based on electromagnetic fields rather than electric fields, and the main measurement equipment is antenna. Antenna, as a sensor, can transmit and receive electromagnetic waves of different frequency bands. Since the electromagnetic wave has characteristics such as refraction and reflection, it is available to analyze the characteristics of the medium in the field by receiving different data. Therefore, non-destructive testing can be achieved based on electromagnetic wave/antenna detection methods.

Detection of multiphase flow by using convolutional neural network (CNN) is inspired by image recognition. One major use of CNN is to identify and classify images. Besides, in recent years, there have been some achievements in dealing with regression issues such as housing prices prediction [5]. The advantage of this method is that no necessary formula derivation is required, as well as that the already trained model can be directly used in the prediction to save time.

2.1. Theory of Antenna Testing

2.1.1. *Maxwell's Equations and Wave Equation*

Maxwell's equations are a set of partial differential equations that the British physicist James Clark Maxwell (1831-1879) established in the 19th century to

describe the relationship between electric fields, magnetic fields, charge density, and current density. This system of equations is composed of four equations: Gauss's law, Gauss's law for magnetism, Faraday's law and Ampère's law with Maxwell's addition. Below is the mathematical expression of the equations [59]:

$$\nabla \times E = -dB/dt$$
$$\nabla \times H = J + dB/dt$$
$$\nabla \cdot D = \rho$$
$$\nabla \cdot B = 0$$

(2.1)

where ρ is the charge density, E is the electric field, B is the magnetic flux density, H is the magnetic field, J is the current density, D is the electric displacement, $\nabla = \frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z}$ is a vector operator.

According to Fourier's theory, many complex situations are composed of linear combinations of harmonic. In the case of a time-harmonic with the time factor $e^{j\omega t}$, the frequency will be single [57]. From Equation (2.1), we have [59]

$$\nabla \times E = -j\omega\mu H$$
$$\nabla \times H = (\sigma + j\omega\mu)E$$
$$\nabla \cdot D = \rho/\varepsilon$$
$$\nabla \cdot B = 0$$

(2.2)

After curling, the first equation can be transformed into:

$$\nabla \times \nabla \times E = \nabla(\nabla \cdot E) - \nabla^2 E = --j\omega\mu\nabla \times H$$

where $\nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$.

Combine this result with the second and third formulas in Equation (2.2) to obtain

$$\nabla^2 E - j\omega\mu(\sigma + j\omega\mu)E = \nabla(\rho/\varepsilon)$$

(2.3)

Define a value γ and let

$$\gamma = \sqrt{j\omega\mu(\sigma + j\omega\mu)} = \alpha + j\beta$$

(2.4)

where α and β are the attenuation constant and phase constant, respectively. These two constants could be derived from Equation (2.4):

$$\alpha = \omega\sqrt{\mu\varepsilon}\left[\frac{1}{2}\left(\sqrt{1 + \frac{\sigma^2}{\varepsilon^2\omega^2}} - 1\right)\right]^{1/2}$$

$$\beta = \omega\sqrt{\mu\varepsilon}\left[\frac{1}{2}\left(\sqrt{1 + \frac{\sigma^2}{\varepsilon^2\omega^2}} + 1\right)\right]^{1/2}$$

Then Equation (2.3) is available to be written as [59]:

$$\nabla^2 E - \gamma^2 E = \nabla(\rho/\varepsilon) \tag{2.5}$$

In the case of source-free region ($\rho = 0$), Equation (2.5) is:

$$\nabla^2 E - \gamma^2 E = 0 \tag{2.6}$$

Here is the wave equation.

2.1.2. Antennas to Radio Waves

Any situation can be considered as a combination of many single frequency sources. If only the case of a single frequency source is considered, Maxwell's equations can be rewritten as [59]:

$$\begin{aligned} \nabla \times E &= -j\omega\mu H \\ \nabla \times H &= J + j\omega\mu E \\ \nabla \cdot D &= \rho/\varepsilon \\ \nabla \cdot B &= 0 \end{aligned} \tag{2.7}$$

Use the derivation process similar to Equation (2.3) to get:

$$\nabla^2 E + \omega^2 \mu\varepsilon E = j\omega\mu J + \nabla(\rho/\varepsilon)$$

This equation links the radiated electric field (no magnetic) to source directly. The boundary conditions are needed to solve the equation. If the boundary is open, the field will disappear when the distance from the source V to the field is infinite [59]. The solution of the Equation (2.7) in a uniformed medium (μ and ε) is [60]

$$E(r) = -j\omega\mu \int_v J(r') \frac{e^{-j\beta \cdot |r-r'|}}{4\pi|r - r'|} dv' + \frac{1}{j\omega\mu} \nabla(\nabla \cdot \int_v J(r') \frac{e^{-j\beta \cdot |r-r'|}}{4\pi|r - r'|} dv')$$

Here r is the distance vector from the origin to the observation point and r' is from the origin to the source point. The radiated electric field of the time varying current J (the time factor $e^{j\omega t}$ is omitted here) is given in this equation which reveals the relationship between radio waves and antennas [59], which is the foundation of antenna theory.

2.1.3. Antenna Bandwidth

As one of the fundamental antenna parameters, Bandwidth indicates the range of frequencies that antenna can exactly release or receive energy. Usually, the required bandwidth is one of the important parameters in the antenna required for decision making. For instance, there are many kinds of antennas with narrow bandwidths which cannot be applied in the wideband operation.

In the experiment, the frequency range in which the antenna can work below -10dB in free space is generally regarded as the bandwidth of the antenna.

2.1.4. Scattering Parameters

Fig. 2.1 shows an ordinary two-port network. a_1 and a_2 are the input whilst b_1 and b_2 are the output at Port 1 and Port 2, respectively. This example can be used to describe scattering parameters:

$$[S] = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{11} \end{bmatrix} \tag{2.8}$$
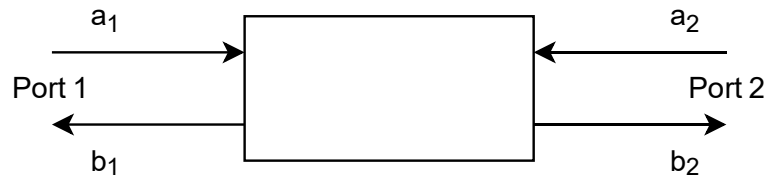
Fig. 2.1 A two-port network

If Equation (2.8) is used to link input with output, then

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{11} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{2.9}$$

Thus,

S_{11} = Port 1 reflection coefficient = b_1/a_1;

S_{12} = Port 2 to Port 1 transmission coefficient/gain = b_1/a_2;

S_{21} = Port 1 to Port 2 transmission coefficient/gain = b_2/a_1;

S_{22} = Port 2 reflection coefficient = b_2/a_2;

In fact, S-parameters are reflection and transmission coefficients not only for a network of 2 ports such as in this case, but also for a network of N ports. There

parameters were initially presented in optics where optical waves were decentralized by objects. Then the concepts were developed to radio waves and RF engineering. However, the term 'S-parameters' remains unchanged.

For a 2-port loss-free network, we have

$$[S_{11}]^2 + [S_{21}]^2 = 1$$
$$[S_{22}]^2 + [S_{12}]^2 = 1 \qquad (2.10)$$

Which is equivalent to

$$|a_1|^2 + |a_2|^2 = |b_1|^2 + |b_2|^2$$

That is, the input power and the output power are the same.

We can know clearly from Fig. 2.2 that a transmitting-receiving antenna system in the space can be regarded as a 2-port network. Furthermore, the transmission and reflection can be characterized by using S-parameters. S_{11} and S_{22} are the reflection coefficients of Antenna 1 and Antenna 2 respectively. They show that the antenna satisfies the antenna feed line well. S_{21} and S_{12} are the transmission coefficients from one antenna to another. They are decided by the features of both antennas, such as radiation patterns and matching, as well as the distance between them.
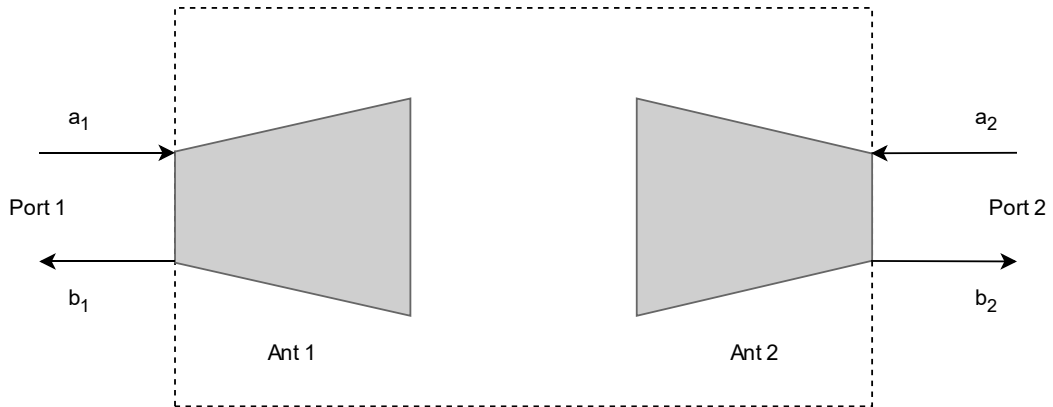
Fig. 2.2 A transmitting-receiving antenna system

2.2. Theory of CNN Model

A convolutional neural network is made up by an input and output layer in combination with different hidden layers. In the hidden layer, there are usually some convolutional layers which convolve with multiplication or other dot product.

These layers are commonly connected to additional convolutions, such as pooling layers or fully connected layers (i.e. dense layers). Their inputs and outputs are hidden by the activation function and the final convolution.

The theory of convolutional neural network is illustrated in Fig. 2.3. The input of CNN model is a tensor with shape (number of samples)×(width)×(height)×(depth). This is an attribute of convolutional layer. To illustrate, a normal picture could be used as an example. The picture size, such as 256×256, means that there are 256×256 pixels which could be describe as numerical values in this picture. Also, each picture has their own depth which is always called RGB of image. Thus, one 256×256 picture could be equal to a 256×256×3 high dimensional matrix.
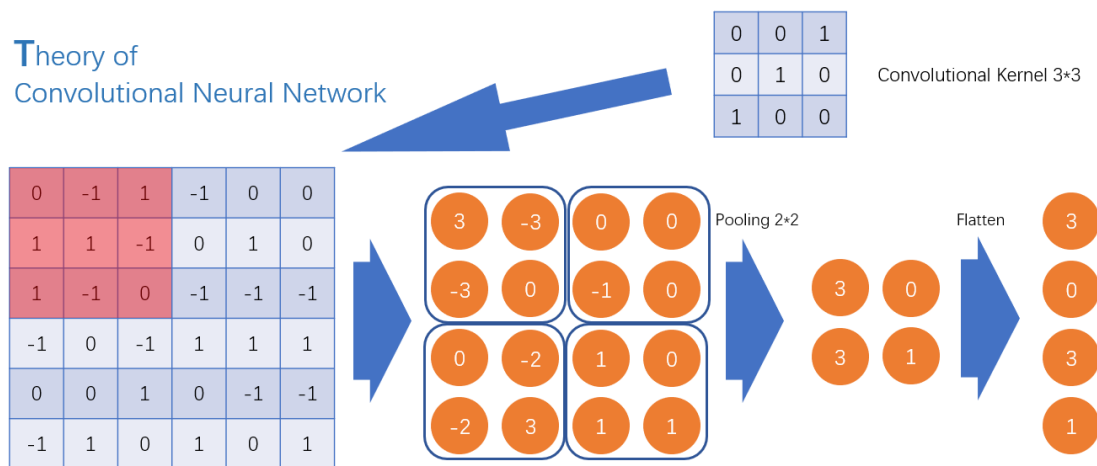
Fig. 2.3 Simple schematic of the convolutional layer

The input in Fig. 2.3, as a simplest example, is a 6×6×1 matrix. It will convolve with a pre-specified 3×3 kernel. The depth of kernel (or filter) is always equal to depth of input. Thus, this kernel here should have the same depth with the input. The multiplication would result in the orange circles which make up a new matrix. The followed layer is pooling layer. There are many kinds of pooling layers, and here a maximum pooling layer is used. It divides all the data in matrix by the same pre-specified size (here is 2×2) and retains only the maximum value in each part. After that, there will be a flattening process to flatten all the left data from the high dimensional matrix (width×height×depth) to a one-dimensional matrix (1×N). Then the fully connected layers are followed. The above process is summarized and drawn as Fig. 2.4, which is the complete model structure diagram.
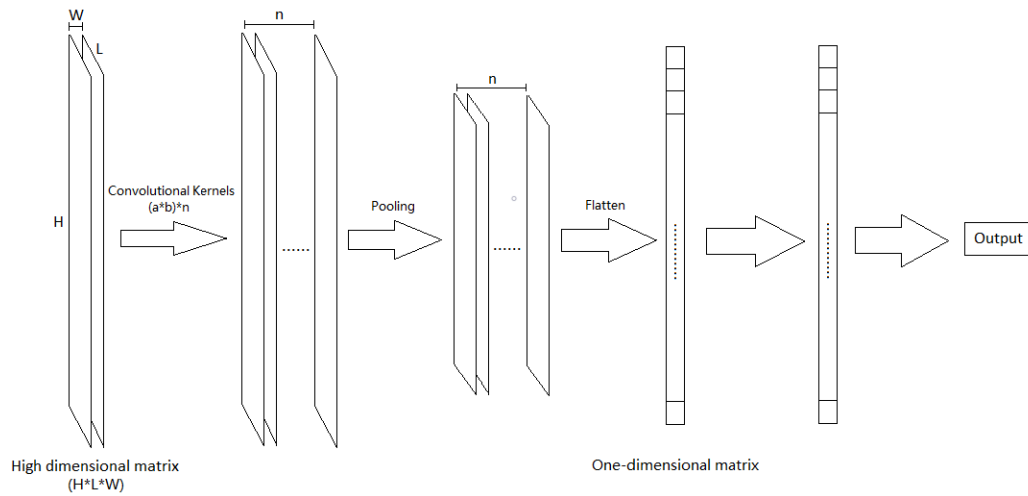
Fig. 2.4 A complete structure of convolutional neural
network with one convolutional layer

In the multiphase flow testing, if the type and content of the liquid in the tube changes, the data detected by the antenna will change accordingly. In other words, the matrix formed by all the parameters and data measured by antennas has a one-to-one correspondence with the environment inside the tube. Analogous to image recognition, although it is difficult to obtain a precise cross-section of the liquid in the tube, each liquid cross-section has a unique matrix corresponding to it. Thus, the liquid content, which could be directly measured by the liquid cross section in the tube, can be indirectly measured by analyzing the matrix.

Fig. 2.5 The shape of
input data.

The length S of the data matrix is the number of parameters, which are always S-parameter. N is the height of matrix, which represents the number of frequency points. The signal, which collected by VNA, has real part and imaginary part. This could be the depth of the data matrix. Therefore, the size of this high dimensional matrix should be $S \times N \times D$ (D=2). Then this data matrix is used as the input of CNN model.

30

CHAPTER 3 MEASUREMENTS

In this chapter, a microwave-based system is employed to obtain the information of a multiphase flow inside a pipe. Unlike a real-world system, a plastic container is used to simulate a pipe. Different amounts of water and oil is filled into the container to simulate different multiphase flow. An Anritsu Vector Network Analyzer (VNA) and two pyramidal horn antennas are used to detect three-phase flow with different liquid content ratios inside the pipe/container. In the measurement, the two horn antennas are used to measure different amounts of liquids in the vertical and horizontal directions. The measured S-parameters are presented as high dimensional matrix before introduced into the convolutional neural network model for training. This neural network model is used to directly predict the water and oil content, thereby avoiding complicated calculation derivation. In data processing, raw data is extracted by VNA, then reshaped the data size into the required high dimensional matrix. The processed data is imported into the neural network model which is written by Python.

3.1.   Measurement Setup
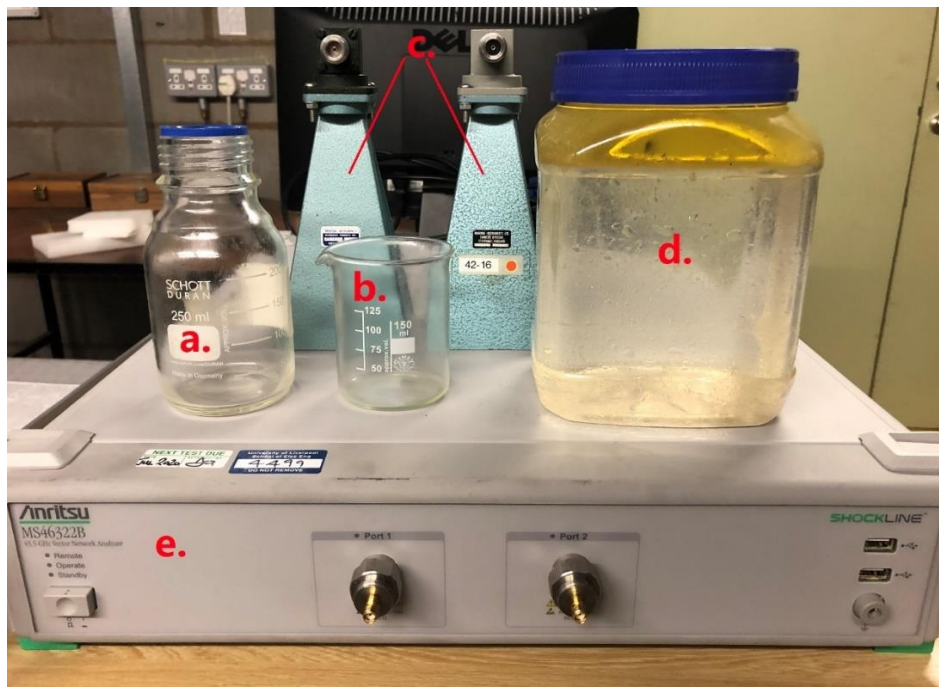
Fig. 3.1 All the devices used in the experiment (a: 250ml flask, b: 150ml beaker, c: two Marconi 6036/4 horn antennas, d: plastic box for simulating pipes, e: two-port Anritsu economy VNA MS46322B).

This experiment uses two horn antennas (Marconi 6036/4) of the same specification as a probing tool for 8 to 12 GHz, a two-port Anritsu VNA MS46322B (frequency range: 1MHz - 43.5GHz) as a transmitting and receiving device, a two-liter (2L) transparent cuboid plastic box to stimulate pipe, which contains vegetable oil and pure water. For accurate measurement of 5% (100 mL) of liquid, this test uses a 250 mL flask and a 200 mL beaker to measure water and oil, respectively. Fig. 3.1 shows all the devices used in this experiment.

For three-phase flow measurement, the permittivity of each phase usually has a large discrepancy. The permittivity of air is considered as 1 (real part of permittivity is 1 and imaginary part of permittivity is 0). After testing, the permittivity of vegetable oil and pure water is shown in the following table.

Table 3.1 Permittivity of vegetable oil and pure water in different frequency.

Frequency, GHz	Permittivity of Vegetable Oil		Permittivity of Pure Water	
	Real Part	Imaginary Part	Real part	Imaginary Part
8	2.32	0.81	66.32	29.13
8.5	2.38	0.70	64.86	30.17
9	2.31	0.58	63.56	31.34
9.5	2.36	0.70	61.92	32.23
10	2.24	0.60	60.61	32.91

As shown in Fig. 3.2, the measurement of this experiment is divided into two parts. The first part is the two antennas placed opposite longitudinally. Then a pipe is placed with a certain proportion of liquid in the middle close to the two antennas. The current directions of the two antennas are parallel to the orientation of the pipe. The S-parameters of this two-port system at this liquid ratio is measured. The second part is to place the two antennas horizontally opposite each other, and then place a pipe with a certain proportion of liquid in the middle close to the two antennas. The current directions of the two antennas are perpendicular to the orientation of the pipe. The S-parameters of this two-port system at this liquid ratio is measured.

During the experiment, in order to facilitate the measurement, the two-phase liquid

(water and oil) needs to remain stationary and layered. Since the two antennas need to remain opposite to each other in a straight line that is either vertical (in vertical testing) or parallel (in horizontal testing) to the horizontal plane, two devices for fixing the two antennas and the pipe need to be fabricated.
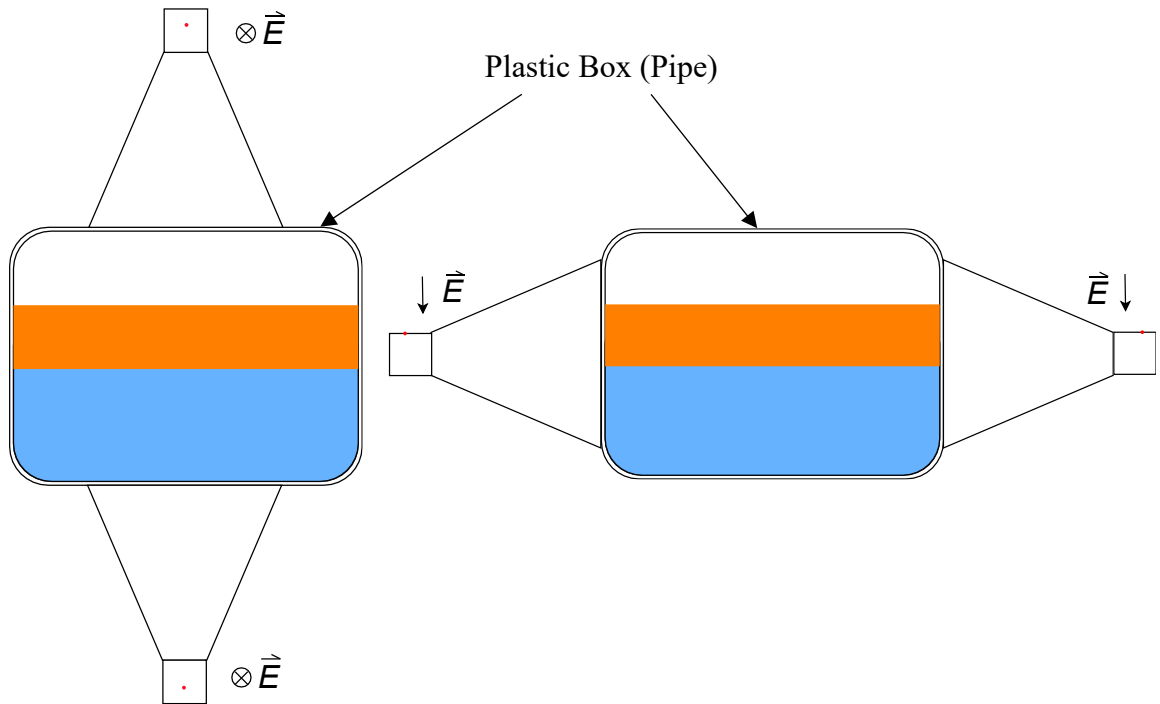
Fig. 3.2 Schematic image of the installation of the two parts in this experiment, the orange part represents oil and the blue part represents water

Fig. 3.3 and Fig. 3.4 show the device pictures for vertical and horizontal testing, respectively, which were built in the laboratory of the High Frequency Engineering Group at the University of Liverpool.

At the time of measurement, there are twenty values to be measured (5% measurement accuracy) from 0% to 95%. By mixing these different contents of water and oil, a total of 210 liquid combinations can be obtained (Section 4.2). Each combination was measured five times in both vertical testing and horizontal testing, with the first four times as training data and the last measured data as test data. A total of 5×210×2=2100 data files of the form '.csv' is obtained. These are the raw data of this experiment.
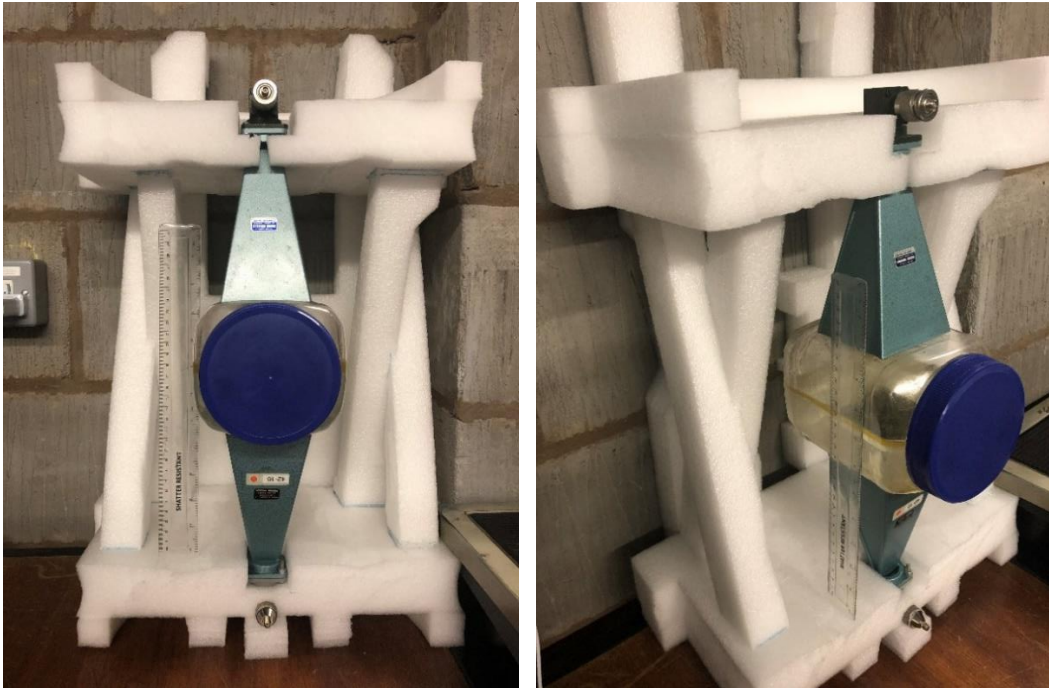
Fig. 3.3 The device pictures for vertical testing
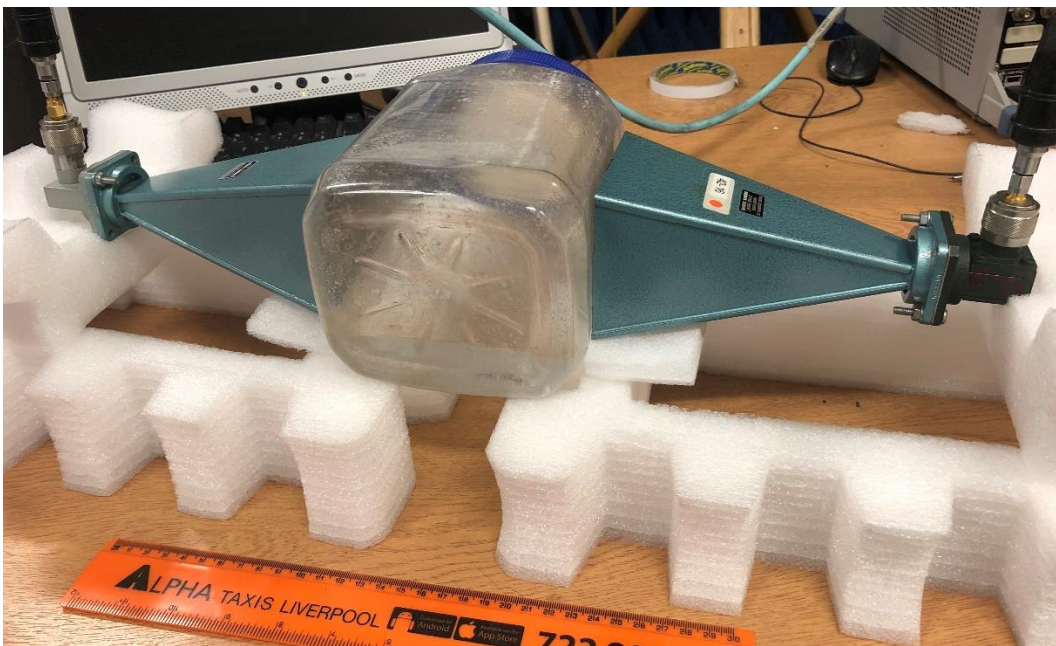
Fig. 3.4 The device picture for horizontal testing

3.1.1. Frequency Chosen

The frequency f and wavelength λ of electromagnetic wave inversely proportional. The relationship between these two parameters follows the formula below:

$$\lambda = \frac{C_0}{\sqrt{\varepsilon_r} \cdot f}$$

34

where C_0 is the speed of light in vacuum; ε_r is the relative permittivity.

The path loss increases as the frequency of the electromagnetic wave increases. Diffraction is more likely to occur after the wavelength becomes longer (the object to be measured is circumvented). Therefore, as the size of the measured object determines the frequency range, a reasonable size is required to prevent excessive path loss which could result from high frequency.

Fig. 3.5 is the plastic box used to simulate the pipe in this experiment. The length H is 125 mm, width W is 110 mm and height H is 150 mm. In order to ensure the measurement sensitivity of the stratified liquid, the wavelength should not be too long. Therefore, the electromagnetic wave propagation path is about 3 to 5 wavelengths here. This ensures the sensitivity of the measurement with an acceptable level of path loss. Considering the above reasons, 8 GHz to 10 GHz is selected as the measured frequency band.
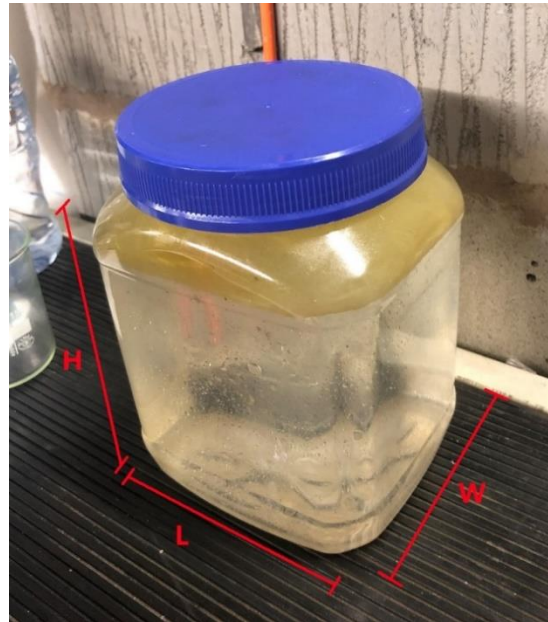
Fig. 3.5 The plastic box used to simulate a pipe

3.1.2.   Antenna

The antennas used in this experiment are Marconi 6036/4 horn antennas. The specifications of this type of antenna are as follows:

Table 3.2 Specifications of Marconi 6036/4 horn antennas

Design	WR90 Standard Gain Horn
Waveguide Size	WR90
Connector type	UBR100 Waveguide Flange
Impedance	50 Ohms
Gain	17.2 dB
Polarization	Linear

Size	Length	80 mm
	Width	80 mm
	Height	150 mm

Fig. 3.6 is the picture of Marconi 6036/4 horn antennas. To ensure that the bandwidth of this type of antenna could cover the frequency range in this experiment (8 GHz – 10 GHz), the testing of bandwidth has been done and the result is shown in Fig. 3.7.

Fig. 3.6 Marconi 6036/4 horn antennas in University of Liverpool
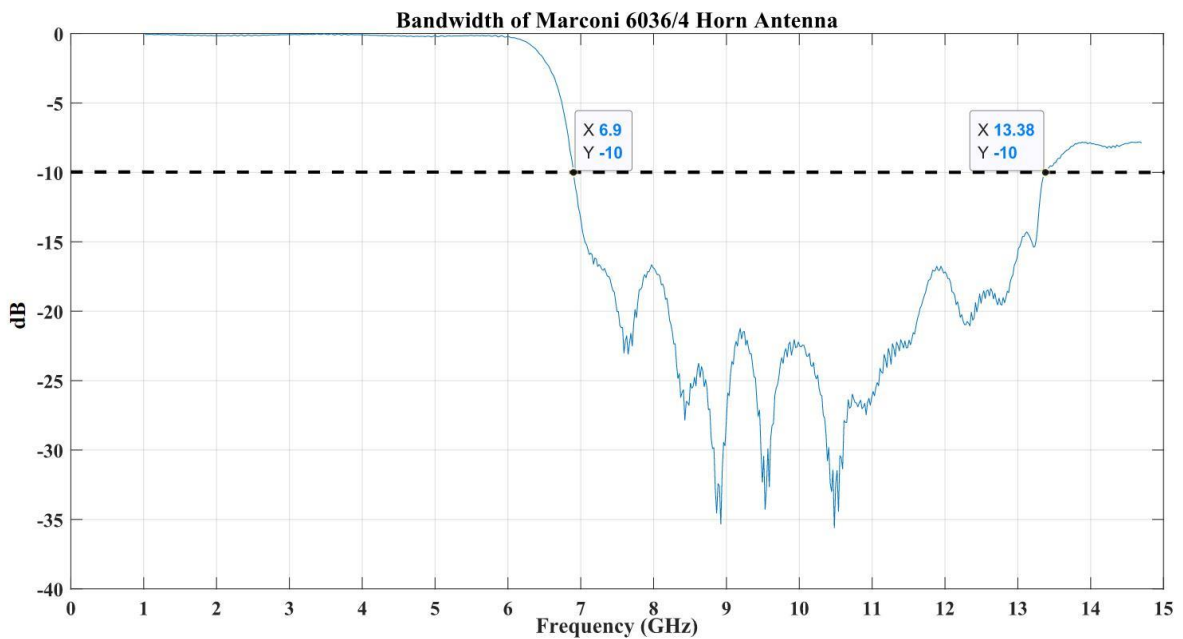
Fig. 3.7 Bandwidth of Marconi 6036/4 horn antenna.

The bandwidth of this type of antenna is from 6.9 GHz to 13.4 GHz as shown in the figure, which covers the frequency range in this experiment.

3.1.3. *Orthogonal Polarization*

If the two 2-port system experiments in Section 3.1 are to be considered as a 4-port system experiment, an experiment is conducted to verify that in these two placement methods, the cross-polarized setup receives much less power than co-polarized setup.
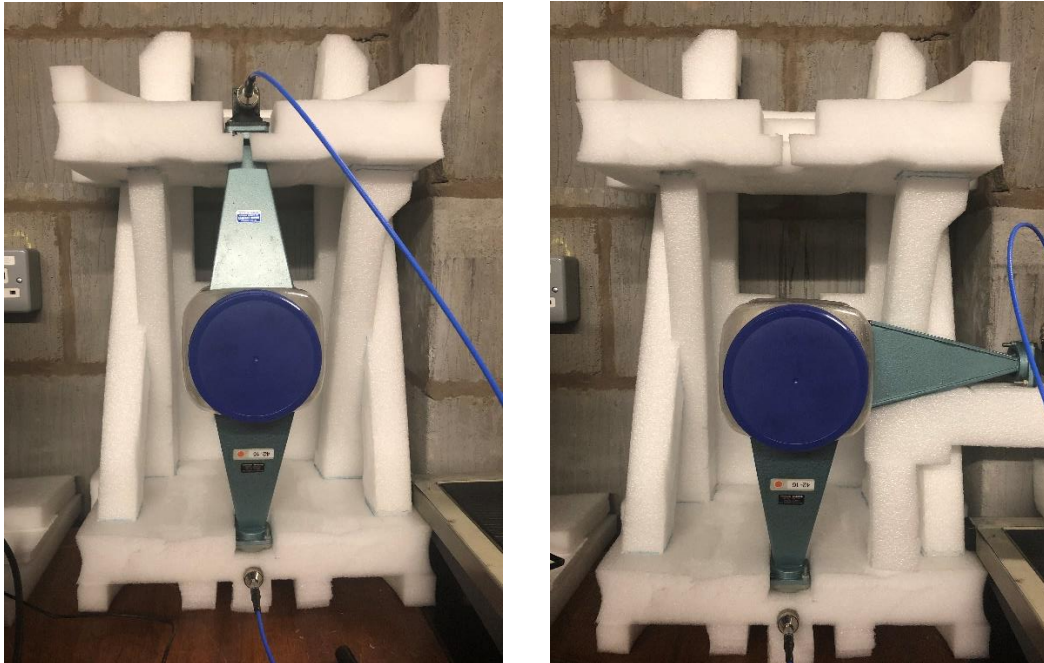
Fig. 3.8 Verification experiment of polarization matching
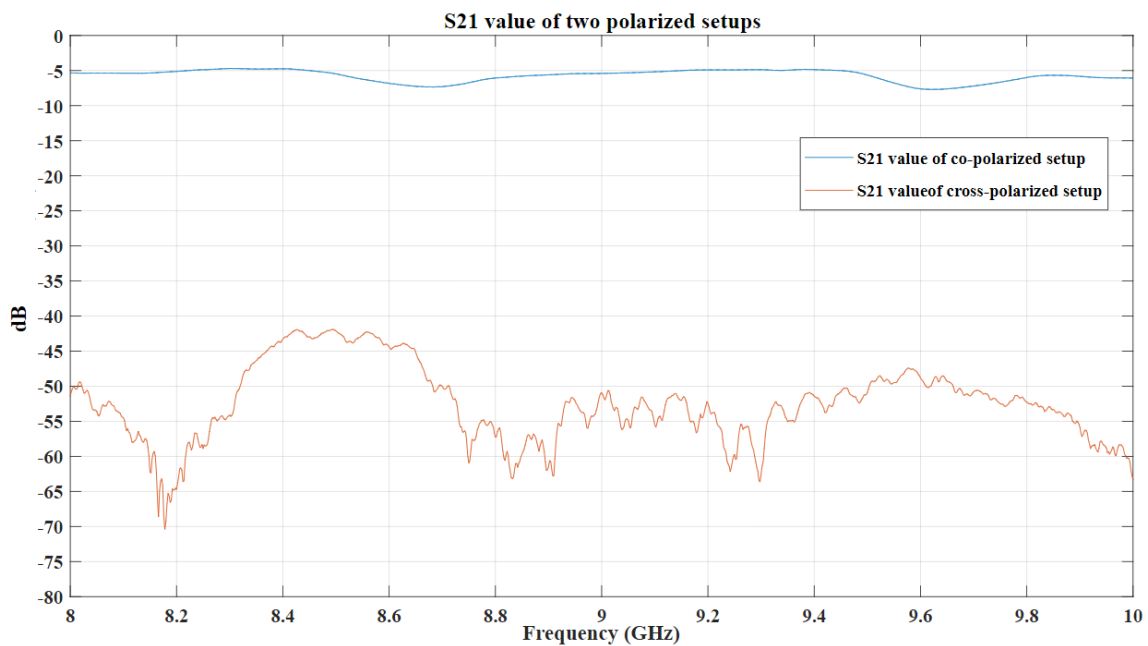
Fig. 3.9 S21 value of two polarized setups

Fig. 3.8 shows the details of the experiment. The antenna is placed in the same way as Fig. 3.2. The S21 value of two polarized setups is shown in Fig. 3.9. The S21 value of cross-polarized setup is 40 dB more than the value of co-polarized setup.

3.2.   Procedure

3.2.1.   *Data Processing*

Table 3.3 shows an example of the raw data recorded by the VNA. There are four parts labeled by number 1, 2, 3 and 4 in this table which correspond to S11, S12, S21 and S22 respectively. Each part has three columns, namely FREQ.GHZ (frequency), REAL (the value of real part) and IMAG (the value of imaginary part). The file exported by the VNA is in '.csv' format. By utilizing MATLAB, the data in these files can be preprocessed.

Table 3.3 An example of the raw data which is recorded by the VNA

PNT	FREQ1. GHZ	REAL 1	IMAG1	FREQ2. GHZ	REAL2	IMAG2	FREQ3. GHZ	REAL3	IMAG3	FREQ4. GHZ	REAL4	IMAG4
1	8	-1.09E-01	9.48E-03	8	2.00E-03	-1.74E-03	8	1.57E-03	-7.30E-04	8	-4.01E-02	-1.89E-01
2	8.002	-1.16E-01	1.18E-02	8.002	1.79E-03	-1.68E-03	8.002	1.73E-03	-5.65E-04	8.002	-5.14E-02	-1.88E-01
3	8.004	-1.15E-01	1.27E-02	8.004	1.79E-03	-1.52E-03	8.004	1.81E-03	-4.28E-04	8.004	-5.60E-02	-1.88E-01
4	8.006	-1.13E-01	1.36E-02	8.006	1.78E-03	-1.40E-03	8.006	1.98E-03	-3.32E-04	8.006	-6.03E-02	-1.88E-01
5	8.008	-1.12E-01	1.48E-02	8.008	1.94E-03	-1.31E-03	8.008	2.20E-03	-2.84E-04	8.008	-6.44E-02	-1.88E-01
6	8.01	-1.10E-01	1.58E-02	8.01	2.07E-03	-1.34E-03	8.01	2.52E-03	-3.19E-04	8.01	-6.86E-02	-1.88E-01
7	8.012	-1.08E-01	1.68E-02	8.012	2.21E-03	-1.47E-03	8.012	2.78E-03	-4.77E-04	8.012	-7.28E-02	-1.87E-01
8	8.014	-1.06E-01	1.79E-02	8.014	2.34E-03	-1.56E-03	8.014	2.99E-03	-7.57E-04	8.014	-7.68E-02	-1.86E-01
9	8.016	-1.04E-01	1.90E-02	8.016	2.47E-03	-1.77E-03	8.016	3.13E-03	-1.16E-03	8.016	-8.04E-02	-1.85E-01
10	8.018	-1.02E-01	2.00E-02	8.018	2.44E-03	-2.11E-03	8.018	3.16E-03	-1.54E-03	8.018	-8.38E-02	-1.84E-01
11	8.02	-1.00E-01	2.09E-02	8.02	2.25E-03	-2.38E-03	8.02	3.10E-03	-1.95E-03	8.02	-8.72E-02	-1.83E-01
12	8.022	-9.87E-02	2.21E-02	8.022	2.16E-03	-2.57E-03	8.022	2.79E-03	-2.23E-03	8.022	-9.04E-02	-1.82E-01
...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...
994	9.986	-1.31E-01	-1.21E-01	9.986	-7.76E-04	-4.23E-04	9.986	-7.33E-04	-1.87E-04	9.986	-1.04E-01	1.29E-01
995	9.988	-1.33E-01	-1.21E-01	9.988	-8.25E-04	-3.04E-04	9.988	-6.91E-04	-8.90E-05	9.988	-1.01E-01	1.32E-01
996	9.99	-1.34E-01	-1.21E-01	9.99	-8.02E-04	-1.99E-04	9.99	-6.15E-04	4.30E-05	9.99	-9.87E-02	1.35E-01
997	9.992	-1.35E-01	-1.21E-01	9.992	-8.00E-04	-5.29E-05	9.992	-5.38E-04	1.76E-04	9.992	-9.59E-02	1.39E-01
998	9.994	-1.35E-01	-1.22E-01	9.994	-7.26E-04	9.62E-05	9.994	-4.13E-04	2.89E-04	9.994	-9.29E-02	1.42E-01
999	9.996	-1.36E-01	-1.22E-01	9.996	-6.22E-04	2.35E-04	9.996	-2.33E-04	3.80E-04	9.996	-8.98E-02	1.45E-01
1000	9.998	-1.37E-01	-1.23E-01	9.998	-4.62E-04	3.32E-04	9.998	-6.26E-05	3.05E-04	9.998	-8.64E-02	1.48E-01
1001	10	-1.38E-01	-1.23E-01	10	-3.57E-04	3.01E-04	10	7.61E-05	2.33E-04	10	-8.29E-02	1.51E-01

In MATLAB, the built-in method 'csvread' is used to read csv files. In order to ensure that information is retained to the maximum extent, the first 1000 data are extracted from the data of 1001 frequency points, which is convenient for the subsequent convolution process using Python. As aforementioned in Section 2.2, the ideal high-dimensional matrix imported into python should be S×N×D, where S is the number of S parameters, N is the number of frequency points, and the value of D is 2, which means two parts of signal (the real part and the imaginary part). Section 3.1.3 mentions one 2-port horizontal testing and one 2-port vertical testing will be used to simulate one 4-port testing. In order to detect the training accuracy of the 4-port testing compared to the 2-port testing, the 2-port measurement data will be imported into the same model for training. However, since the number of S-parameters of the data (2 sets of 2-port S-parameters), which requires an 8×1000×2 high dimensional matrix for the raw data. In other words, the model's input size is set to 8. Accordingly, a set of 2-port measured data also needs to be processed into a matrix of size 8×1000×2. Therefore, when processing the experimental data of horizontal/vertical testing only, the original data is processed into a matrix of size 4×1000×2. Repeat this step to process another identical matrix. The two matrices of size 4×1000×2 are then spliced into a matrix of size 8×1000×2 and imported into the model. The shape of three input matrices (2-port system of horizontal testing only, 2-port system of vertical testing only, 4-port system of horizontal and vertical testing) are shown in Fig. 3.106.
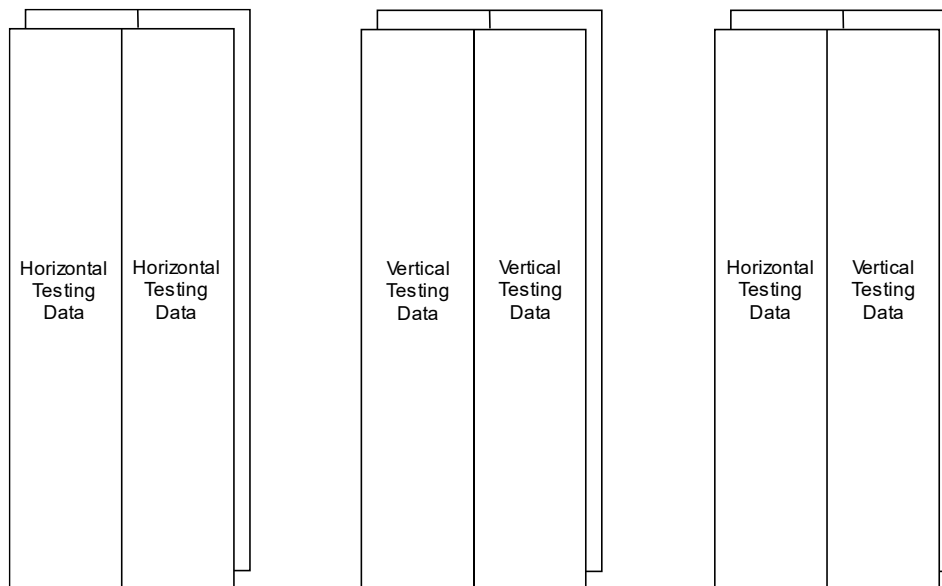
Fig. 3.10 The shape of three input matrices

When importing MATLAB processed data into Python, the 'reshape' operation in Python is used. In Python, since 'reshape' is a reverse fill, for the 8×1000×2 high-dimensional matrix, 'reshape' is filled from '2' first, then filled with 1000×2, and finally an $8 \times 1000 \times 2$ high-dimensional matrix is formed after repeating this process 8 times. Therefore, when processing raw data, the data has been processed in MATLAB as 2×1000×8 (or 2×8000) format and exported as a '.mat' file. Then we use 'sio.loudmat' in Python to read the exported '.mat' files. The figure below shows the data processing flow chart

```
┌─────────────┐
│   Victor    │
│   Network   │
│   Analyzer  │
└─────────────┘
       │
       │  Raw Data
       │  .csv format
       ▼
┌─────────────┐
│   MATLAB    │
│             │
└─────────────┘
       │
       │  2*1000*8
       │  .mat format
       ▼
┌─────────────┐
│Python Reshape│
│             │
└─────────────┘
       │
       │  8*1000*2
       ▼
┌─────────────┐
│  CNN Model  │
│             │
└─────────────┘
```
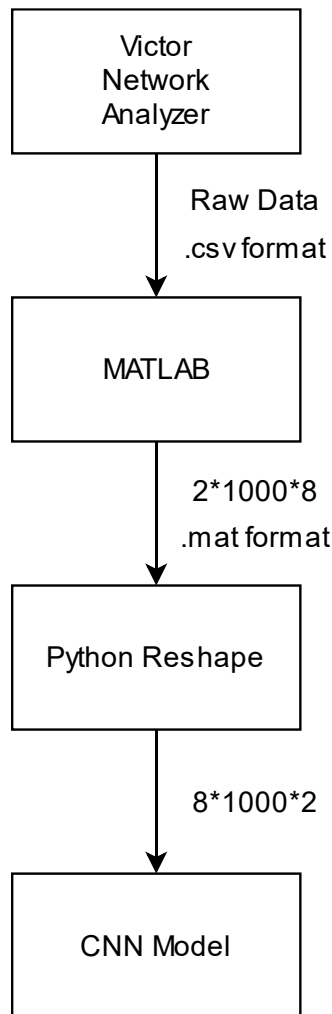
Fig. 3.11 Data processing flow chart

For the target matrix that needs to be imported into Python, its size should be the number of samples multiplied by the number of outputs (water and oil content). For more convenient automatic entry of data, the files' names can be saved as 'water+A+oil+B' when exporting data from the VNA, where A and B are the

volume percentages of water and oil, respectively. Then in MATLAB, regular expressions are used to extract the values of A and B, which are subsequently recorded in the target matrix.

3.2.2 Neural Network Model Design

The convolutional neural network (CNN) model used in this experiment is designed on Keras. Keras is an open source neural network library written in Python. It supports running on top of TensorFlow, Microsoft Cognitive Toolkit, Theano or PlaidML. Keras is developed to achieve rapid experimentation of deep neural networks with the advantages of being user-friendly, modular and scalable.

The output is a matrix of size 2×1, which are water content and oil content. Three convolutional layers and two dense layers are used in this model. Mean squared error (MSE) is chosen as the loss function metric. The network structure is demonstrated in Fig. 3.12.

This network structure figure has the following description:

1. A ReLU operation is performed after each convolutional layer to improve the nonlinear characteristics of the network.

2. The main purpose of the 1×1 convolutional layer is to reduce the dimension as well as to correct the linear activation (ReLU in this case). For example, the output size of the previous layer is 100×100×128. After 5×5 convolutional layer with 256 channels (stride size = 1, padding size = 2), the output data size is $100 \times 100 \times 256$, and the number of convolution layer parameters is 128×5×5×256=819200. If the output of the above layer passes through the 1×1 convolutional layer with 32 channels and then the 5×5 convolutional layer with 256 outputs, the output size is still $100 \times 100 \times 256$, but the convolution parameter has been reduced to $128 \times 1 \times 1 \times 32 +$ 32×5×5×256=204800, a dramatic reduction of about 4 times.
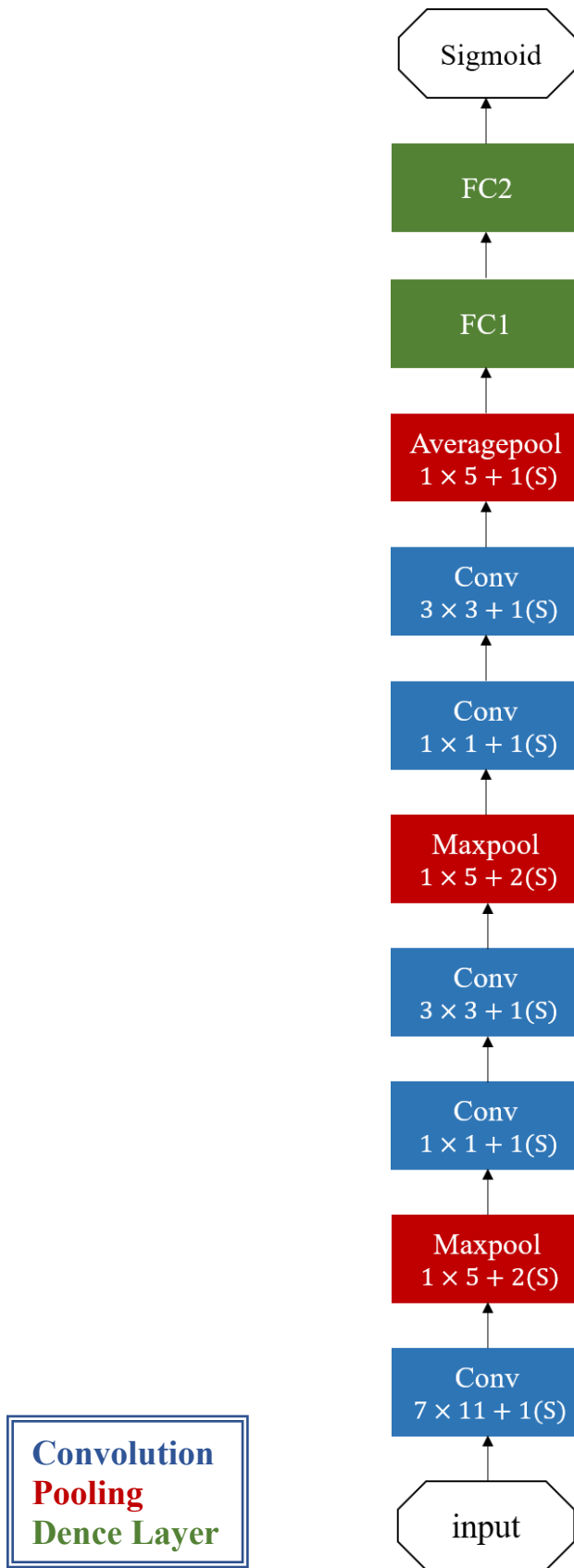
Fig. 3.12 Network model structure

The network structure details are as follows:

1. Input:

   The original input matrix is 8×1000×2.

2. The first layer (convolutional layer):

   Use 7×11 convolution kernel (stride is 1) and 64 channels, output is 8×1000×64. Perform ReLU operation after convolution. After 1×5 max pooling (stride is 1×2), the output is 8×500×64.

3. The second layer (convolutional layer):

   Use 64 convolution kernels with size 1×1 to reduce the dimensionality before the 3×3 convolution kernel and perform ReLU operation. Patch size becomes 8×500×64. Then the result is convolved by 192 3×3 convolution kernels (padding is 1) and output size is 8×500×192. Perform ReLU operation after convolution. After 1×5 max pooling (stride is 1×2), the output size is 8×250×192.

4. The third layer (convolutional layer):

   Use 96 convolution kernels with size 1×1 to reduce the dimensionality before the 3×3 convolution kernel and perform ReLU operation. Patch size becomes 8×250×96. Then the result is convolved by 128 3×3 convolution kernels (padding is 1) and output size is 8×250×128. Perform ReLU operation after convolution. After 1×5 average pooling (stride is 1×1), the output size is 8×250×128.

Table 3.4 The network structure details.

Type	Patch Size / Stride	Output Size	Depth	#3×3 reduce	#3×3	Params
Convolution	7×11 / 1	8×1000×64	1	-	-	9920
Max pool	1×5 / 2	8×500×64	0	-	-	-
Convolution	-	8×500×192	2	64	192	110784
Max pool	1×5 / 2	8×250×192	0	-	-	-
Convolution	-	8×250×128	2	96	128	110720
Average pool	1×5 / 1	8×250×128	0	-	-	-
Flatten	-	1×1×256000	-	-	-	-
Linear	-	1×1×256	1	-	-	65536k
Linear	-	1×1×2	1	-	-	-

CHAPTER 4 RESULTS AND DISCUSSION

4.1. Model Training Result

In machine learning, gradient descent is an iterative optimization algorithm that looks for the best result (the minimum of the curve). To find a local minimum of a function using gradient descent, one takes steps proportional to the negative of the gradient (or approximate gradient) of the function at the current point. This algorithm is iterative, meaning that it has to be applied multiple times to get the optimization results. The iterative nature of the gradient descent allows the graphical representation of the under-fitting to evolve in order to obtain the best fit to the data.

When the data size is too large to be input at one time, the data could be divided and input into computer as small pieces one by one. The number of data in these small pieces is called batch size. These batches update the weight of the neural network at the end of each step and fit the given data. During training, a batch of training image is trained once through the network (one forward propagation + one backward propagation). The weight is updated once per iteration. At the time of testing, one batch of test images passes through the network once (one forward propagation). When a completed data set passes through the neural network once and returns once, the process is called one epoch. The number of batches required to complete one epoch is called iterations.

For the training data, a validation split value is set to divide all the training data to two part: training data and validation data. Training data is used to train the model as the name implies, and validation data is used to select and modify the model. After training, testing data is needed to show how the expected condition of this model will be under real conditions. The biggest difference between testing data and validation data is that the model or parameters are modified by validation data only. In this model, the value of batch size is 64 and the value of epoch is set to 300. 15% of training data is randomly selected as validation data. In this neural network model, mean squared error (MSE) is chosen as loss function. After testing, the training and validation error reach convergence by training 300 epochs for this convolutional neural network model. Fig. 4.1 illustrates the train and validation

error by training 300 epochs in Python.
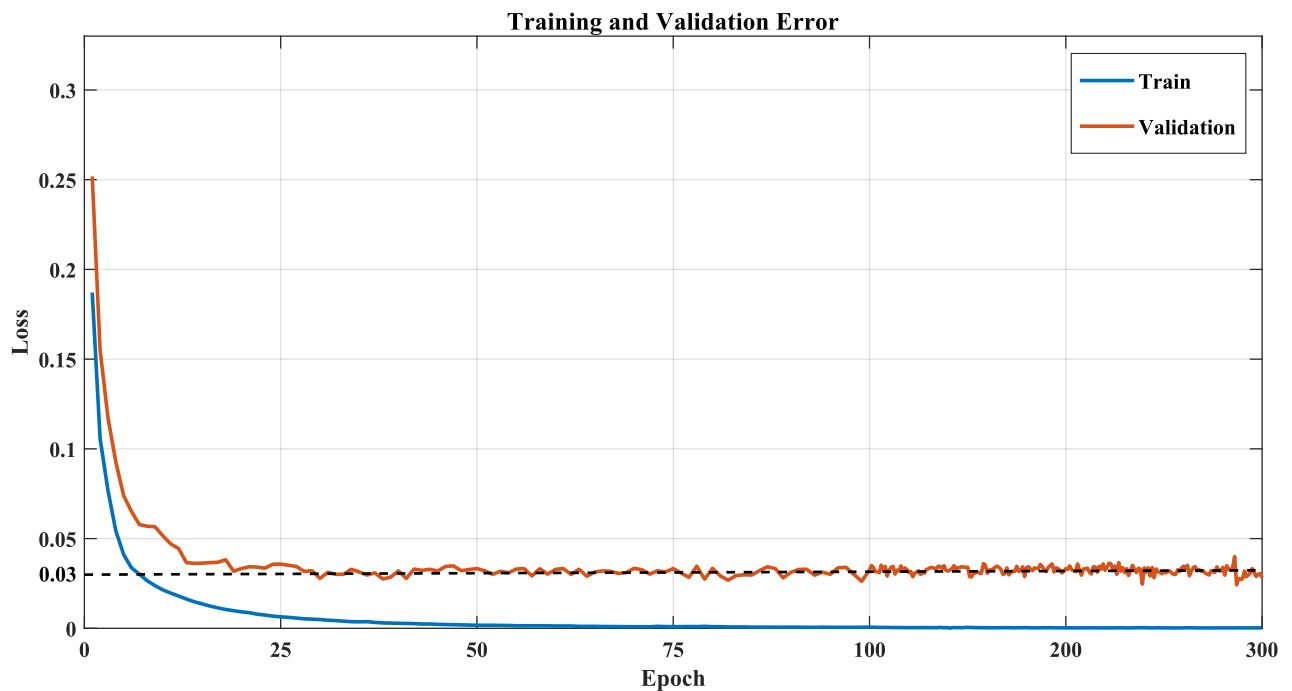
Fig. 4.1 Training and validation error for 300 epochs

As can be seen in the figure, the validation loss is higher than the training loss, indicating that there is data overfitting in the model training. This situation is common when using neural network models for regression analysis. The training accuracy of 300 epochs in the figure is acceptable in this experiment.

4.2.　Horizontal Testing and Vertical Testing

It is inaccurate to combine the data of horizontal testing and vertical testing as the output data of the 4-port system. But due to the characteristics of this experiment, only the content of different phase is changed. No change or movement of the antennas and the pipe occurred during the measurement. In other words, whether it is the data of the 4-port system or the splicing data of the two 2-ports, the influence of the environment outside the tube and the mutual influence between the antennas are constant during a complete experiment. It is also demonstrated in Section 3.1.3 that there is no significant influence between the antenna placed laterally and the antenna placed longitudinally in terms of orthogonal polarization (the effect on the signal between the opposing antennas is not significant). Thus, this data combination was built to simulate the measurement data of the 4-port antenna

system and compare the training results with the results of the 2-port system after the model training.

As mentioned in Section 3.1, there are a total of 210 combinations of different three-phase liquid contents in the measurement. Table 4.1 and Table 4.2 show all combinations and list the prediction of oil content and water content for each combination.

Table 4.1 The prediction of oil content tested by the neural network which use horizontal data and vertical data as input.

	The true content of oil in the pipeline, %																			
	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
0	1.2	6.4	9.3	14.3	20.9	26.2	30.2	35.8	39.1	44.6	49.8	55.9	61.8	66.0	69.3	74.6	80.2	85.7	90.0	94.4
5	0.3	4.1	10.2	15.2	19.3	25.9	30.3	36.1	40.8	45.9	51.9	55.4	59.6	65.6	69.8	74.7	80.8	84.8	88.7	
10	0.4	5.0	10.5	15.9	19.7	25.7	30.0	35.6	39.6	46.5	51.0	55.3	61.0	65.1	69.7	73.9	80.2	84.6		
15	0.2	5.4	9.9	15.8	20.2	25.8	30.6	34.0	40.5	45.6	49.7	55.5	61.1	65.3	70.1	75.5	79.9			
20	0.7	5.4	9.3	16.5	19.6	25.3	31.0	35.6	40.7	45.6	51.2	55.9	60.2	65.2	70.0	74.5				
25	0.3	5.6	9.5	15.6	20.4	25.8	29.3	34.9	39.2	44.6	49.8	55.1	60.7	65.0	70.8					
30	0.2	4.4	10.0	15.0	19.5	25.6	29.5	33.8	39.6	44.6	49.2	54.8	59.7	64.9						
35	0.1	4.9	9.0	14.7	19.3	25.0	28.6	34.6	40.5	45.0	49.8	54.7	60.8							
40	0.1	4.2	9.6	13.8	19.6	23.6	28.2	34.6	39.0	44.8	49.7	54.3								
45	0.1	4.6	10.1	15.5	19.7	24.2	28.5	34.5	38.8	44.4	50.4									
50	0.2	3.6	8.6	15.3	19.3	24.2	28.9	34.6	39.3	44.7										
55	0.2	4.8	9.1	15.8	19.3	25.3	28.7	34.7	40.0											
60	0.0	4.6	8.9	14.7	19.0	23.1	31.6	36.7												
65	0.3	3.9	8.7	13.3	20.9	26.3	32.0													
70	0.1	6.4	9.6	13.1	22.7	26.7														
75	0.1	5.9	9.6	17.9	21.9															
80	1.3	7.6	8.7	15.2																
85	0.2	6.0	11.0																	
90	0.1	9.1																		
95	0.1																			

The true content of water in the pipeline, %

The predicted value of the oil and water are separately compiled and analyzed for better analysis result. It is easy to see from Table 4.1 that almost all the prediction results of oil have little error compared with the theoretical value. A relatively large error in the prediction of 5% oil only occurs when the water content reaches 90%. Fig.4.2 and Fig. 4.3 are comparison diagrams based on the data in Table 4.1 which are more intuitive than the comparison of the predicted and theoretical values.
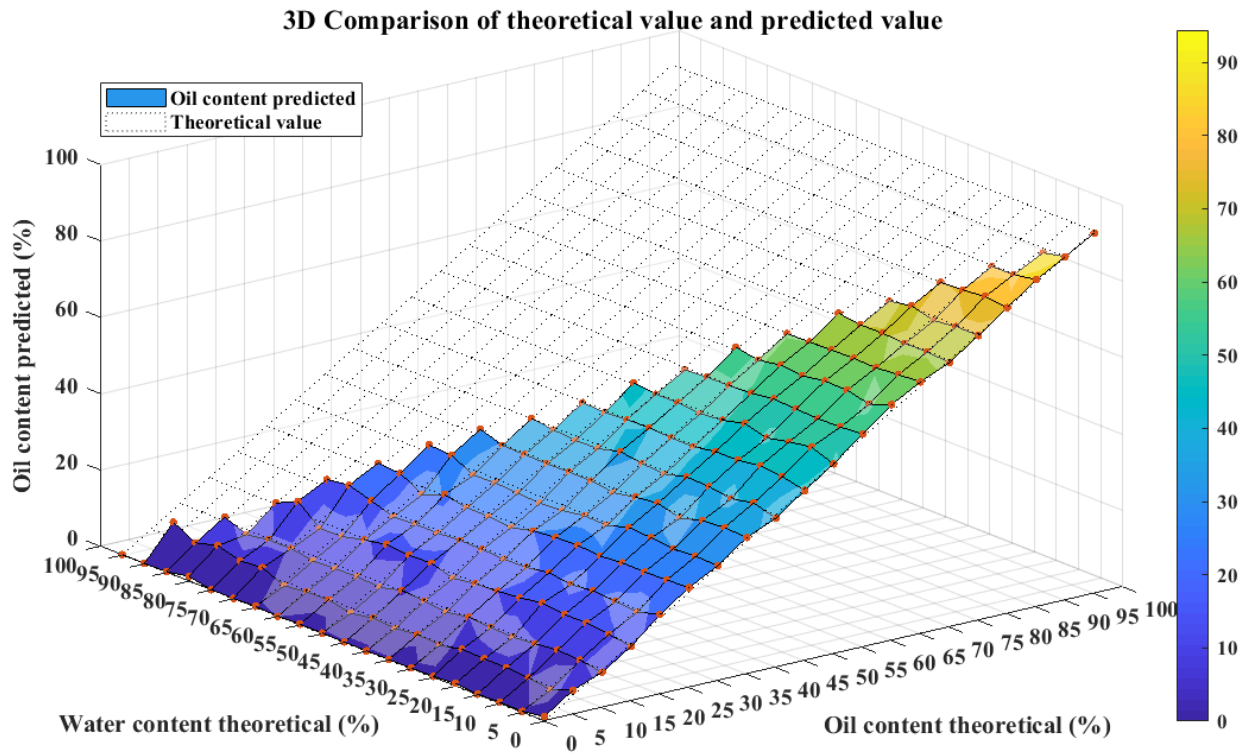
Fig. 4.2 3D Comparison of theoretical value and predicted value (Table 4.1).
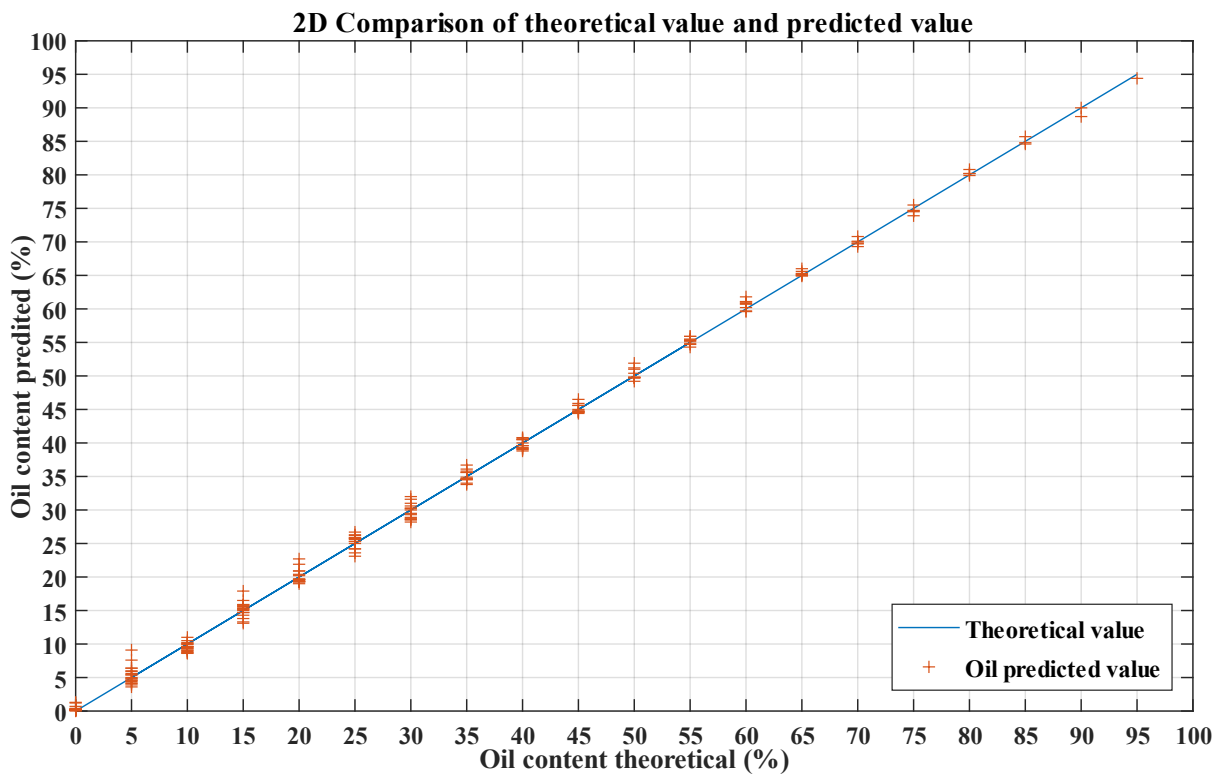
Fig. 4.3 2D Comparison of theoretical value and predicted value (Table 4.1).

In 3D Comparison, the closer the predicted image is to the standard value, the smaller the error. In order to calculate the error value, here we use Error Sum of Square (SSE) as the figure of merit.

$$SSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

The following table shows the SSE value in each content of oil.

Table 4.2 The SSE value in each content of oil based on Table 4.1.

Theoretical value, %	0	5	10	15	20	25	30	35	40	45
SSE	0.216	1.86	0.632	1.302	1.027	1.066	1.367	0.672	0.528	0.432
Theoretical value, %	50	55	60	65	70	75	80	85	90	95
SSE	0.715	0.306	0.859	0.216	0.212	0.392	0.183	0.23	0.845	0.36

The SSE value of all the data in Table 4.1 is 0.8072. The water content predicted by neural network model is discussed below.

Table 4.3 The prediction of water content tested by the neural network which use horizontal data and vertical data as input.

|		The true content of water in the pipeline, %																		
		0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
	0	1.0	5.7	10.0	14.6	18.6	24.3	29.8	34.9	39.6	44.2	50.5	55.9	60.7	58.2	59.1	55.8	61.8	57.8	57.3	50.4
	5	0.2	6.1	9.8	14.4	20.0	23.6	29.1	34.7	39.0	44.5	49.5	55.4	59.1	61.9	61.6	63.6	62.1	66.7	64.2	
	10	0.0	5.7	9.1	14.2	18.7	23.5	28.2	34.2	38.5	44.9	50.7	55.3	60.9	57.1	64.1	65.2	63.8	64.3		
	15	0.1	4.9	9.8	14.3	18.7	23.1	29.3	33.6	39.7	44.4	50.5	55.5	60.3	60.2	59.1	58.6	61.7			
	20	0.1	4.5	9.0	14.3	19.4	24.4	29.4	34.2	39.6	44.0	49.9	55.9	60.6	59.7	56.9	59.9				
The true content of oil in the pipeline, %	25	0.1	4.9	8.9	13.8	19.0	24.7	28.5	34.6	39.6	44.2	51.0	55.1	60.2	60.4	59.6					
	30	0.3	4.6	9.3	13.5	19.1	24.3	28.9	35.0	39.9	45.1	50.4	54.8	58.6	62.1						
	35	0.1	4.6	9.5	14.3	18.5	24.2	28.8	34.4	39.6	45.2	50.7	54.7	58.2							
	40	0.1	4.8	9.4	13.8	19.3	23.8	28.6	34.2	39.7	45.3	51.1	54.3								
	45	0.3	4.8	9.8	13.6	19.2	24.3	29.0	34.9	40.0	44.9	50.7									
	50	0.2	4.6	10.3	13.4	19.3	24.4	28.4	34.2	39.9	44.4										
	55	0.2	4.7	8.6	14.5	18.8	24.0	29.0	34.3	40.5											
	60	0.2	4.5	9.3	14.0	19.0	24.0	29.1	34.3												
	65	0.2	4.6	9.8	14.1	18.9	24.3	29.0													
	70	0.2	4.8	9.7	13.5	19.2	24.0														
	75	0.4	4.4	8.9	13.5	19.3															
	80	0.3	4.1	9.0	15.0																
	85	0.3	4.3	10.0																	
	90	0.3	3.9																		
	95	0.2																			

Fig. 4.4 3D Comparison of theoretical value and predicted value (Table 4.3).
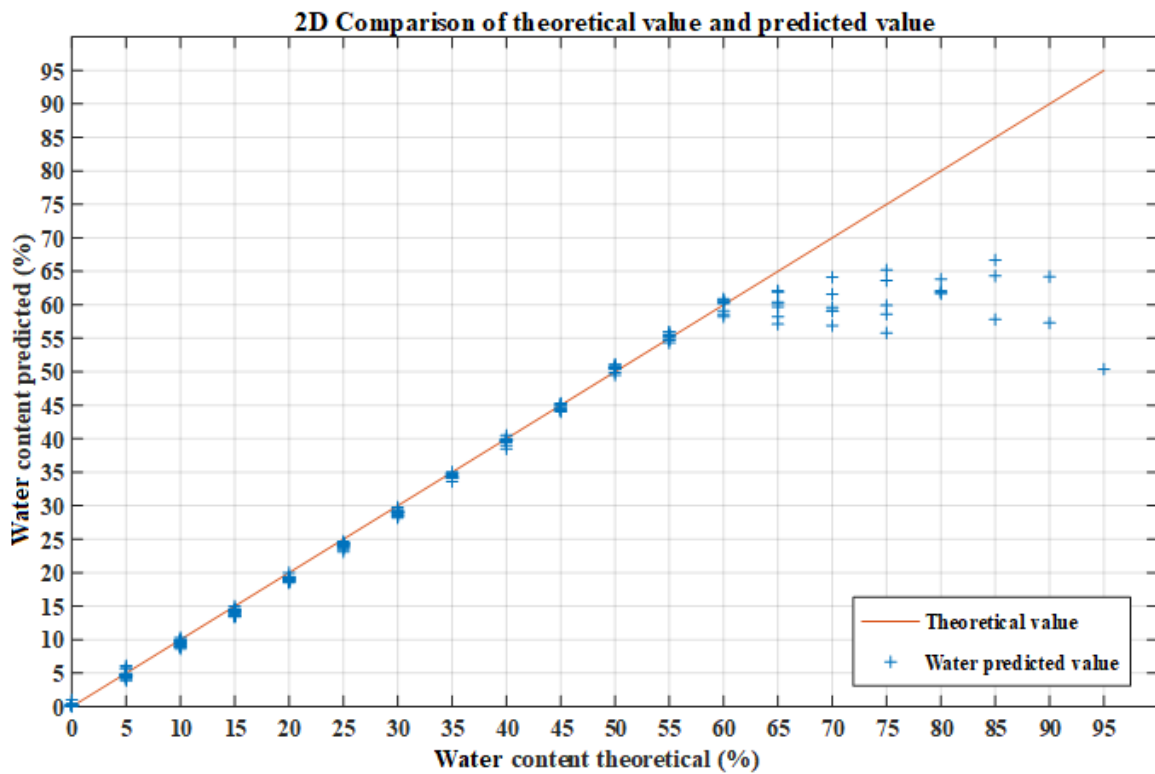
Fig. 4.5 2D Comparison of theoretical value and predicted value (Table 4.3).

It can be seen from the above two figures that the model is more accurate when the predicted water content is below 60%. When it is higher than 60%, the predicted result is no longer accurate. A table of the SSE values in each content of water is shown below.

Table 4.4 The SSE values in each content of water based on Table 4.3.

Theoretical value, %	0	5	10	15	20	25	30	35	40	45
SSE	0.097	0.338	0.507	1.108	1.01	1.045	1.298	0.472	0.362	0.31
Theoretical value, %	50	55	60	65	70	75	80	85	90	95
SSE	0.46	0.306	0.975	28.42	103.8	218.3	312.2	501.1	867.5	1989.1

And the SSE value of all the data in Table 4.3 is 40.5.

4.3. Horizontal Testing Only

The purpose of training and testing the data of the 2-port system separately is to test the effect of increasing the number of antennas on the accuracy of the predicted results.

Table 4.5 The prediction of oil content tested by the neural network which only use horizontal data as input.

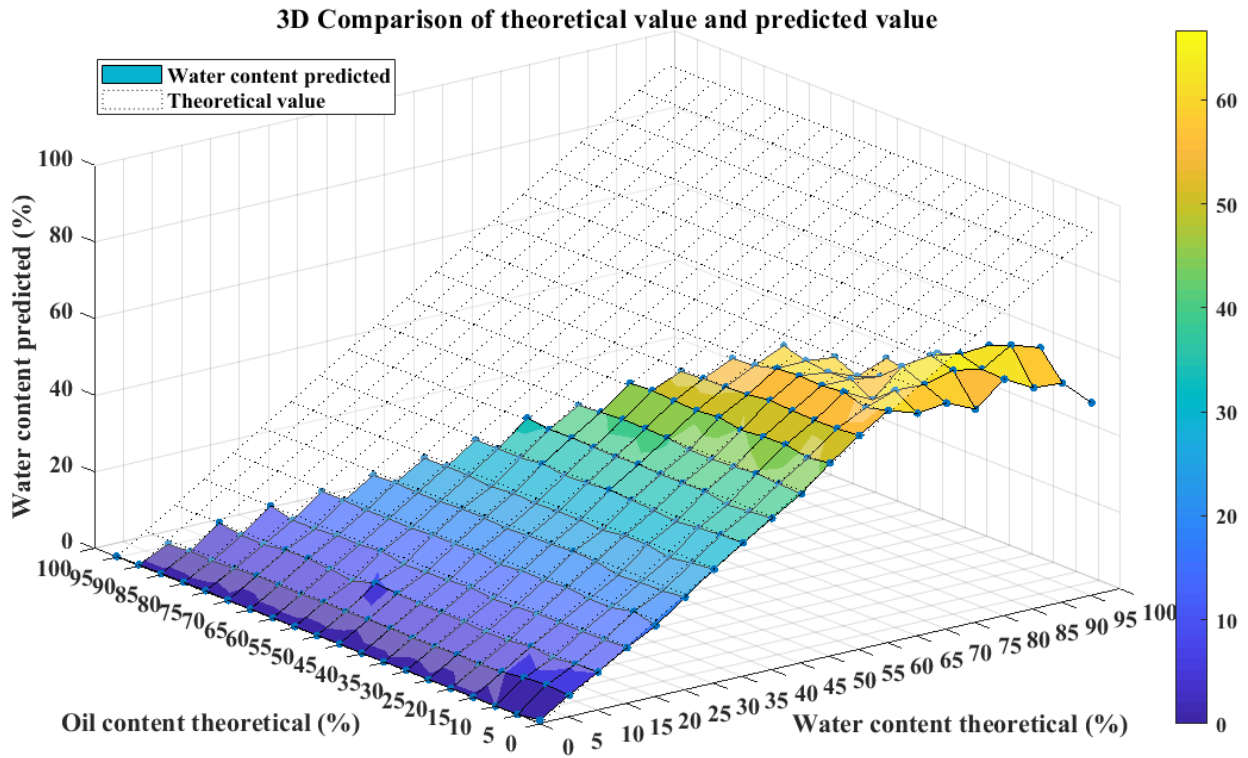|		The true content of oil in the pipeline, %																		
		0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
The true content of water in the pipeline, %	0	1.4	5.7	7.1	13.6	18.5	24.5	13.9	33.7	38.9	43.9	48.8	54.4	61.0	65.7	68.7	75.0	80.1	85.5	90.1	92.3
	5	0.5	4.8	8.9	12.5	18.5	24.2	29.3	33.8	38.6	43.8	48.3	56.4	59.0	63.7	69.6	78.0	82.1	83.5	88.1	
	10	0.7	5.0	8.4	29.5	21.2	25.5	29.3	34.7	38.8	43.3	50.4	53.6	58.0	64.1	67.4	72.7	80.0	72.0		
	15	0.3	15.3	9.0	14.9	20.1	23.8	28.7	30.5	40.4	44.2	50.5	55.7	59.3	64.8	70.0	74.1	73.6			
	20	1.3	5.2	10.3	15.3	31.8	25.3	29.2	37.1	39.1	43.2	49.3	52.9	57.5	66.5	69.4	73.6				
	25	0.4	5.3	11.1	10.7	12.1	24.5	30.3	34.2	38.8	42.7	48.5	47.8	60.0	65.0	60.6					
	30	0.3	20.4	9.2	14.7	21.0	25.2	36.7	33.6	40.0	44.3	21.5	58.5	60.0	64.4						
	35	0.1	4.1	10.1	15.5	3.7	32.0	28.8	40.2	39.2	45.3	48.1	54.7	59.6							
	40	0.0	10.9	9.6	14.0	19.6	25.6	29.0	35.0	40.7	44.4	50.0	49.0								
	45	0.1	3.2	11.0	28.6	19.8	24.8	41.6	33.7	38.5	37.1	40.6									
	50	0.2	11.7	12.5	14.9	19.1	24.4	25.0	21.8	38.7	42.7										
	55	0.2	9.0	9.5	15.2	15.4	23.8	28.3	31.4	27.9											
	60	0.2	0.4	8.7	15.6	17.9	21.8	22.8	27.0												
	65	4.1	0.5	11.5	20.3	25.0	26.1	29.5													
	70	1.1	5.4	12.4	27.0	31.0	31.1														
	75	1.2	5.3	12.2	31.4	34.2															
	80	1.1	10.3	14.9	39.0																
	85	2.1	11.4	13.6																	
	90	1.4	13.5																		
	95	1.4																			

Fig. 4.6 3D Comparison of theoretical value and predicted value (Table 4.5).
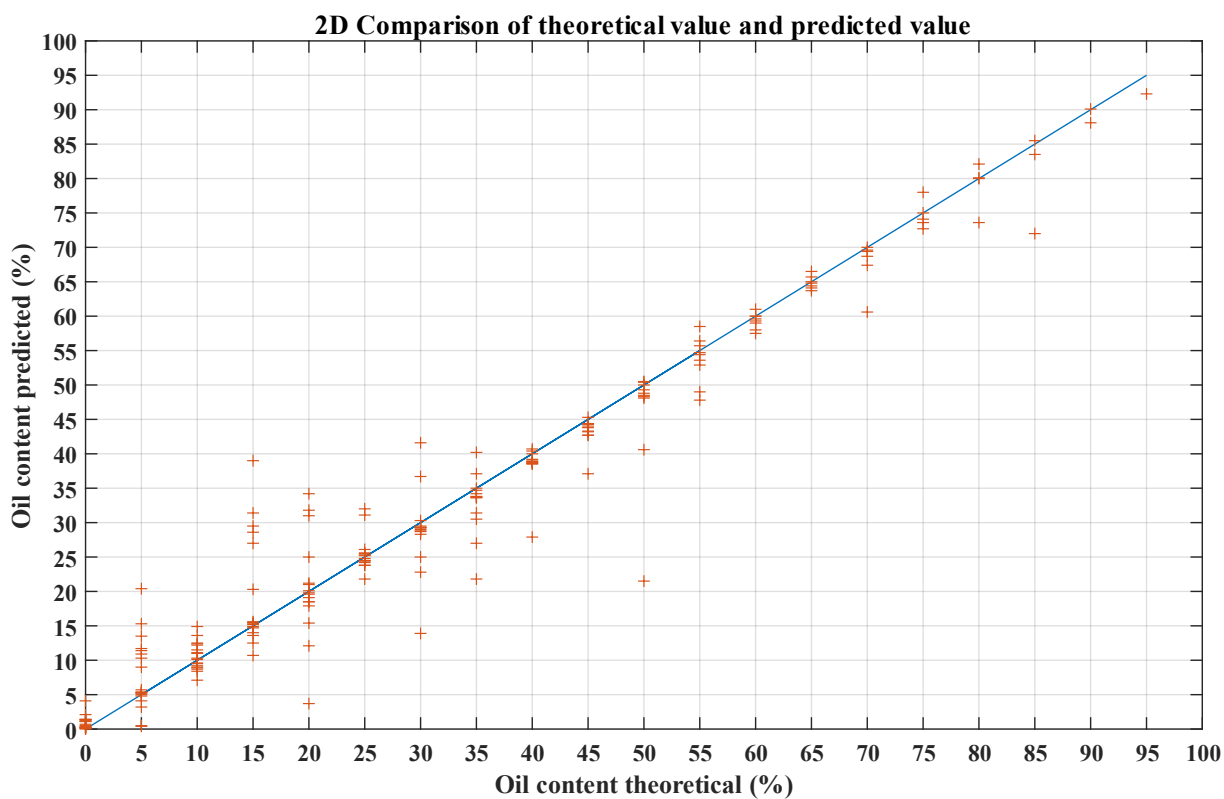
Fig. 4.7 2D Comparison of theoretical value and predicted value (Table 4.5).

Table 4.5 illustrates the prediction of oil content tested by the neural network which only uses horizontal data as input. The error of the data can be clearly seen from Figures 2 and 3. Therefore, if only horizontal testing data are used for training and detection, a successful prediction could still be achieved, though the accuracy is not high enough. The SSE value of the data are shown as follows.

Table 4.6 The SSE values in each content of water based on Table 4.5.

Theoretical value, %	0	5	10	15	20	25	30	35	40	45
SSE	1.694	32.98	4.129	84.75	53.03	6.855	37.46	23.88	13.21	7.577
Theoretical value, %	50	55	60	65	70	75	80	85	90	95
SSE	91.17	12.15	1.613	0.806	16.22	3.412	11.35	57.17	1.81	7.29

The SSE value of all the data in Table 4.3 is 26.6.

The predicted results of water content are below.

Table 4.7 The prediction of water content tested by the neural network which only use horizontal data as input.

	The true content of water in the pipeline, %																			
	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
0	2.5	5.0	11.1	15.5	20.1	26.3	29.1	35.0	39.8	44.0	50.0	55.2	58.5	16.0	28.6	17.2	25.4	54.9	28.4	27.6
5	4.4	6.7	10.0	10.5	20.1	24.8	19.2	24.8	15.1	44.6	11.8	4.4	6.4	16.3	61.4	64.6	63.6	63.6	62.0	
10	2.3	2.1	8.9	14.5	20.0	25.0	30.0	34.3	40.6	43.5	46.6	53.0	58.7	62.3	64.0	65.1	65.7	66.1		
15	1.3	4.1	10.4	15.4	20.0	27.7	30.3	34.6	40.0	40.9	50.2	53.4	56.5	60.1	59.5	58.3	53.3			
20	2.2	6.8	9.6	14.0	21.1	28.7	29.7	25.7	38.7	44.1	48.8	52.6	59.1	59.7	60.8	56.8				
25	0.9	5.6	11.0	16.3	20.0	24.6	28.7	33.9	35.9	43.9	48.4	53.5	57.7	59.1	58.3					
30	6.3	5.1	10.5	15.8	21.3	25.7	27.7	34.9	38.6	44.3	45.4	51.6	59.1	57.3						
35	1.4	5.5	9.7	17.0	19.9	25.5	30.0	40.3	38.6	43.5	45.5	52.8	58.7							
40	0.7	4.8	11.0	14.1	20.5	25.0	28.3	33.9	41.7	43.2	48.3	48.3								
45	1.7	4.2	12.0	14.7	19.7	23.3	29.4	34.3	39.1	42.4	48.6									
50	1.0	5.8	9.8	15.7	19.2	24.5	49.4	34.5	39.0	48.6										
55	1.1	4.4	10.5	19.7	17.2	24.0	28.9	35.0	38.2											
60	1.0	5.1	10.1	13.4	20.3	24.0	28.4	34.2												
65	0.9	5.5	10.1	13.5	21.6	25.4	28.6													
70	0.7	5.0	11.6	13.4	20.0	26.4														
75	0.7	4.9	11.6	17.3	20.1															
80	0.6	4.5	11.6	11.9																
85	0.6	5.5	9.6																	
90	0.5	4.7																		
95	0.5																			

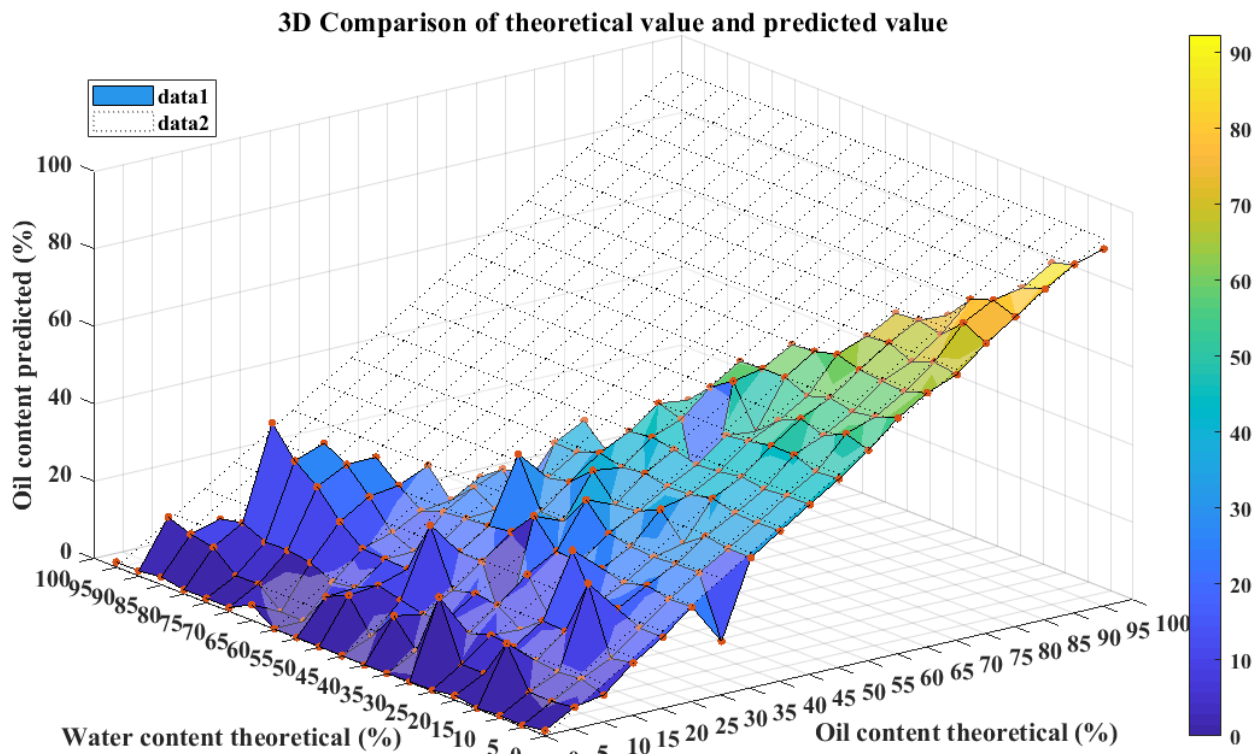The true content of oil in the pipeline, %

Fig. 4.8 3D Comparison of theoretical value and predicted value (Table 4.7).

Fig. 4.9 2D Comparison of theoretical value and predicted value (Table 4.7).

Fig. 4.8 and 4.9 show the predicted results of water content. Results are similar to the water predicted results in Section 4.2 to some extent, but also with lower accuracy.

The SSE value of the data are shown in Table 4.7 below. The SSE value of all the data in Table 4.7 is 166.57.

Table 4.8 The SSE values in each content of water based on Table 4.7.

Theoretical value, %	0	5	10	15	20	25	30	35	40	45
SSE	4.475	0.974	0.957	4.352	0.901	2.058	36.43	17.16	54.23	4.358
Theoretical value, %	50	55	60	65	70	75	80	85	90	95
SSE	152.1	292.9	362.2	703.7	359.3	831.4	1042	573.7	2289	4543

4.4. Vertical Testing Only

For the same purpose in the last section, the experimental data of the vertical testing is also imported into the model for training and testing to check the prediction results.

Table 4.9 The prediction of oil content tested by the neural network which only use vertical data as input.

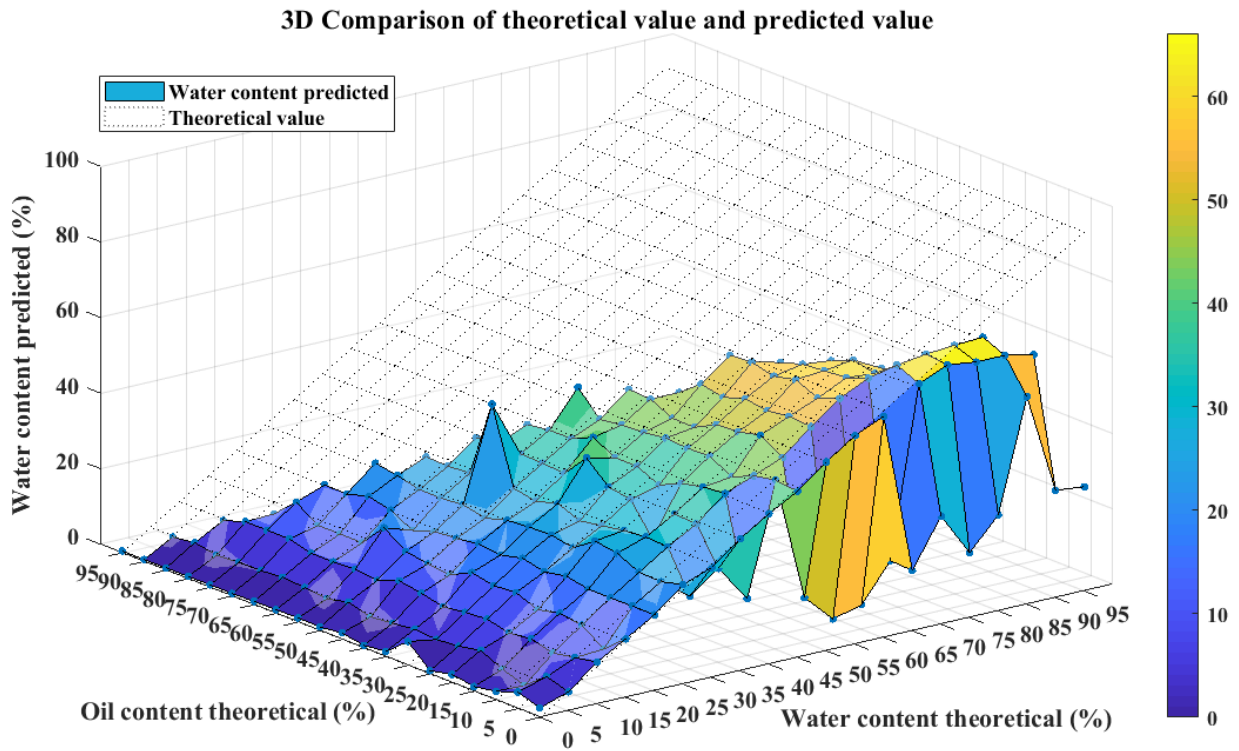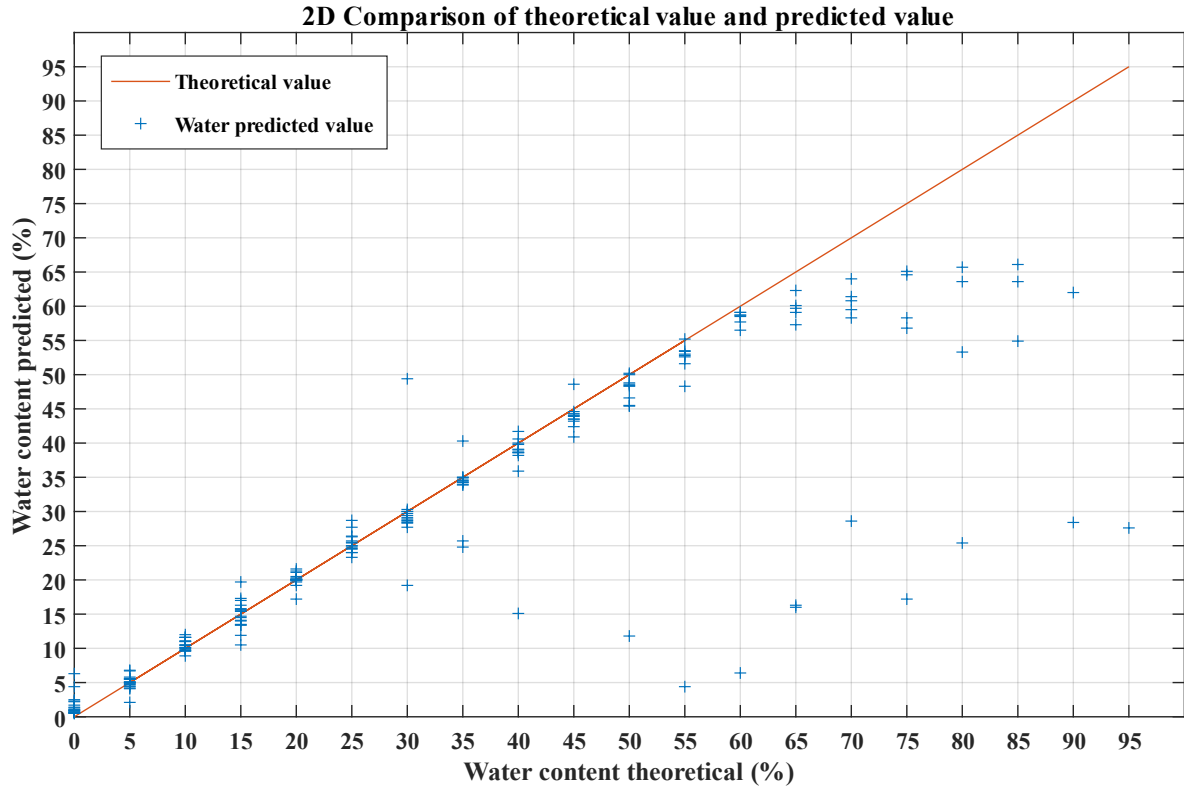| The true content of water in the pipeline, %	The true content of oil in the pipeline, %																			
	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
0	34.8	10.0	16.3	29.0	15.8	21.7	11.5	51.6	42	56.9	43.0	40.3	35.2	57.3	62.5	60.2	51.1	54.6	60.2	62.7
5	5.4	6.2	8.1	13.7	14.9	14.5	11.5	16.5	11.3	25.5	32.8	16.4	11.9	15.0	9.1	24.4	10.3	36.2	18.3	
10	5.6	10.8	7.8	25.6	12.2	19.2	21.2	19.8	27.1	14.6	11.8	21.7	36.7	22.9	23.5	17.6	14.1	42.1		
15	1.2	21	11.7	19.3	36.7	17.1	12.7	13.2	27.2	21.2	27.5	13.1	10.3	18.1	39.6	27	27.3			
20	2.1	3.7	2.5	18.8	19.8	17.2	37.1	9.4	10.1	20.3	32.3	25.1	23.5	17.9	10.6	32.0				
25	3.7	2.7	3.6	16.2	10.3	19.2	25.0	22.7	27.1	8.2	8.2	26.5	31.4	28.5	21.5					
30	1.4	3.0	13.8	20	16.6	10.2	8.4	21.5	26.7	25.8	20.5	25.3	7.9	25.4						
35	0.9	7.1	3.9	17.1	25.7	19.1	22.9	7.1	8.3	22.6	23.8	16.3	13.3							
40	4.3	3.0	3.9	9.8	11.3	21.7	22.1	10.2	17.2	15.9	9.3	24.2								
45	1.8	7.2	7.2	12.8	13.4	10.2	13.8	17.9	21.2	10.8	8.7									
50	2.8	1.9	3.3	18.6	9.3	8.5	11.9	24.5	11.5	25.2										
55	3.6	3.3	5.2	9.8	8.5	14.9	13.6	5.3	9.7											
60	1.5	1.4	5.4	7.1	9.1	17.5	10.7	7.9												
65	8.8	2.2	1.5	11.0	5.5	3.6	14.9													
70	17.9	4.4	3.2	15.6	5.0	5.2														
75	1.1	4.9	5.5	9.0	11.0															
80	4.3	5.9	1.4	5.0																
85	13.2	2.5	1.0																	
90	0.7	7.9																		
95	0.6																			

Fig. 4.10 3D Comparison of theoretical value and predicted value (Table 4.9).

Fig. 4.11 2D Comparison of theoretical value and predicted value (Table 4.9).

Different from the results in Section 4.2 and 4.3, the predicted result of oil content with vertical data only as input is unsatisfying. Table 4.9 illustrates the prediction result and the error of data based on Figures 2 and 3. Results showed that it is difficult to predict the oil content by training the model with vertical testing data only. Table 4.10 shows the SSE value of the data in Table 4.9.

Table 4.10 The SSE values in each content of water based on Table 4.9.

Theoretical value, %	0	5	10	15	20	25	30	35	40	45
SSE	96.56	20.5	34.87	38.85	95.16	138.4	226.8	440.4	500.3	659.6
Theoretical value, %	50	55	60	65	70	75	80	85	90	95
SSE	929	1067	1623	1664	2122	2045	3203	1715	3014	1043

The SSE value of all the data is 575.67 in this part.

Below is the prediction of water content.

Table 4.11 The prediction of oil content tested by the neural network which only use vertical data as input.

	The true content of water in the pipeline, %																			
	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95
0	1.2	34.4	30.2	38.5	52.3	43.8	35.7	53.5	40.4	38.8	52.6	56.4	52.6	28.7	49.1	59.6	42.6	26.1	63.1	60.1
5	10.1	13.3	14.9	6.6	56.5	56.6	49.4	21.9	35.5	24	36.1	50.6	54.1	50.3	44.6	51.6	50.5	57.0	40.1	
10	1.1	11.9	23	28.9	19.4	46.8	42.9	31.6	36.8	51.5	48	37.5	54.8	59.4	49.6	52.2	62.7	61.2		
15	3	6.1	24.6	14.2	17.7	40.2	35.7	26.1	37.4	50.3	42.6	19.3	50.3	55.3	39.7	38.4	57.2			
20	1.3	7.9	26	32.3	30.8	24.8	41.7	41.6	26	32.9	50	50.6	22.1	53.2	51.5	47.6				
25	6.0	4.9	18.7	30.5	18.4	20.0	42.8	43.1	29.9	38.9	53.9	51.1	27.3	58.4	55.5					
30	1.9	6.1	20.4	18.9	17.8	31.9	22.1	27.3	42.2	24.5	24	49.9	47.8	26.2						
35	4.1	14.3	18.3	14.2	43.1	39.3	27.7	45.1	51	41.4	18.9	54.9	53.8							
40	2.1	7.4	6.2	27	21.6	10.7	29.9	28.4	20.9	18.6	26.1	28.6								
45	3.5	6.3	36.9	17.4	16.6	46.1	43.5	40.5	33	52.3	51.7									
50	3.4	10.1	24.0	9.0	11.6	33.3	21.7	13.2	32	43.5										
55	3.1	12.9	11.6	40.0	29.6	21.4	23.4	42.4	38.6											
60	2.7	4.9	16.3	29.2	17.5	7.7	33.8	41.5												
65	3	12.6	21.8	11.2	19.4	28.2	35.2													
70	3.2	23.9	12.6	17.8	32.5	41.0														
75	4.8	9.2	11.5	24.0	22.4															
80	3.2	5.1	33.6	30.8																
85	4.0	11.0	28.0																	
90	3.2	21.4																		
95	4.2																			

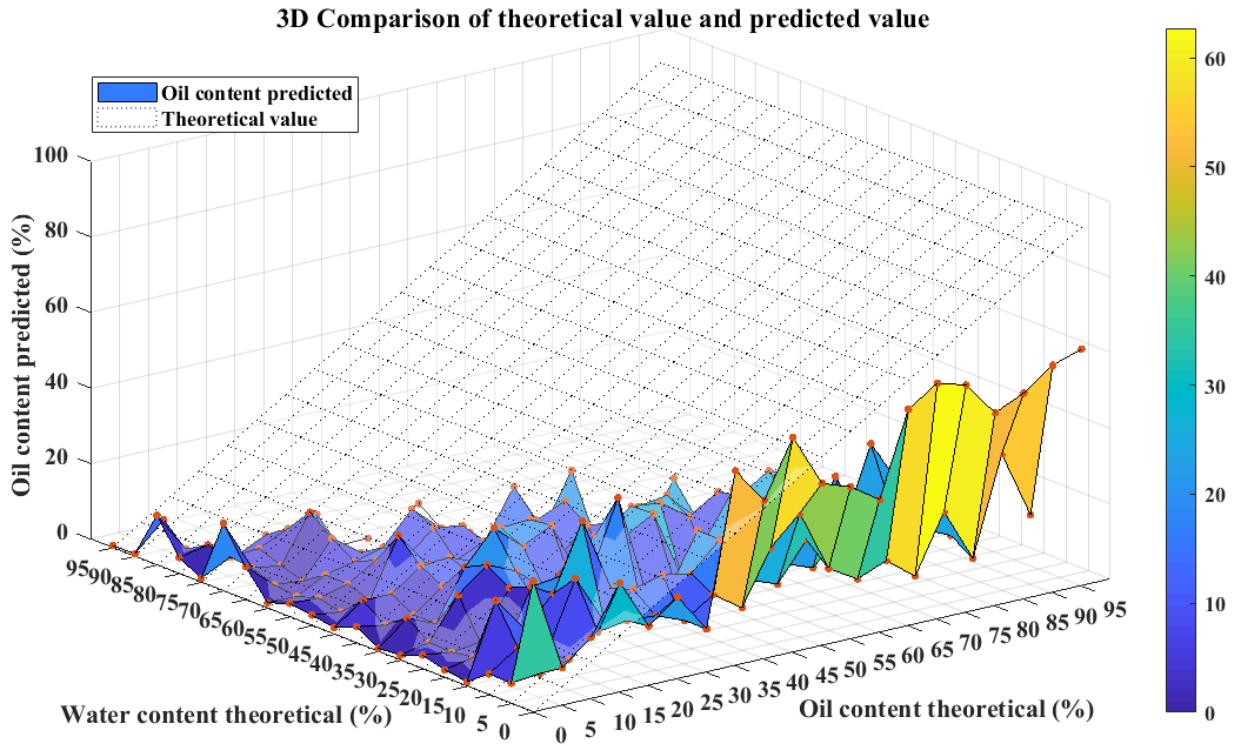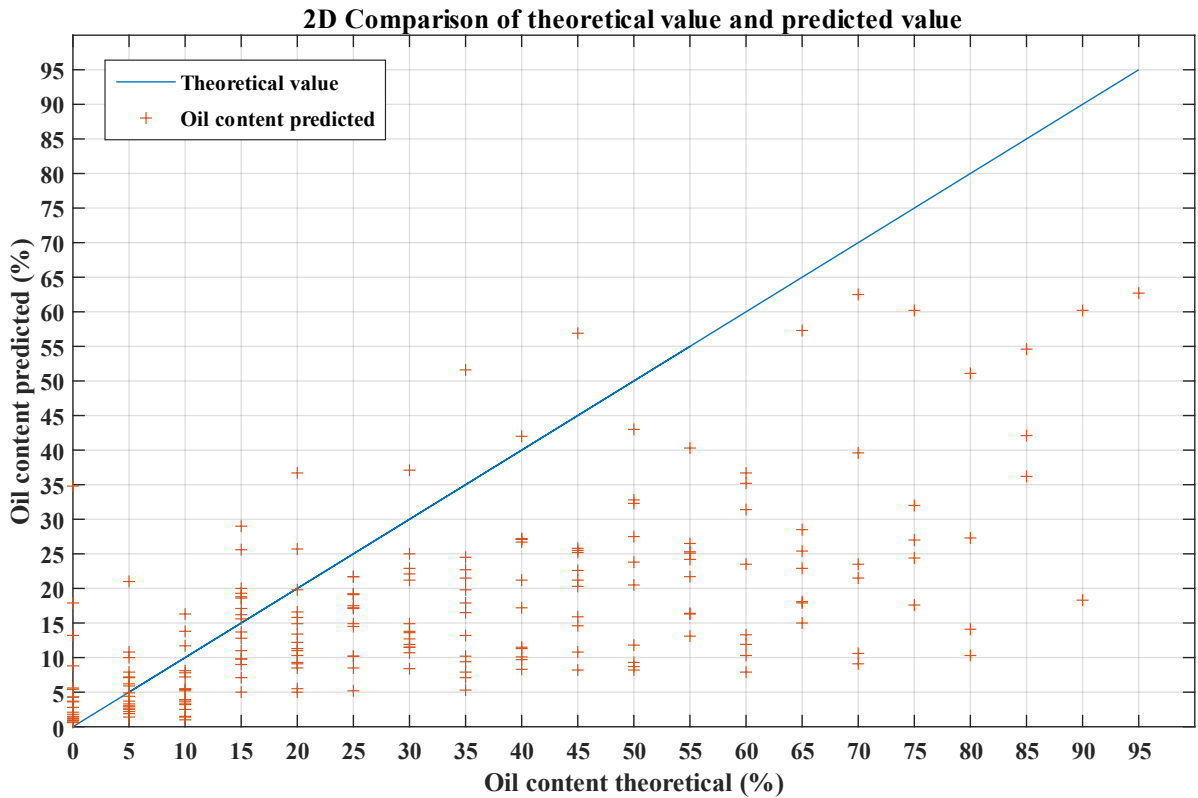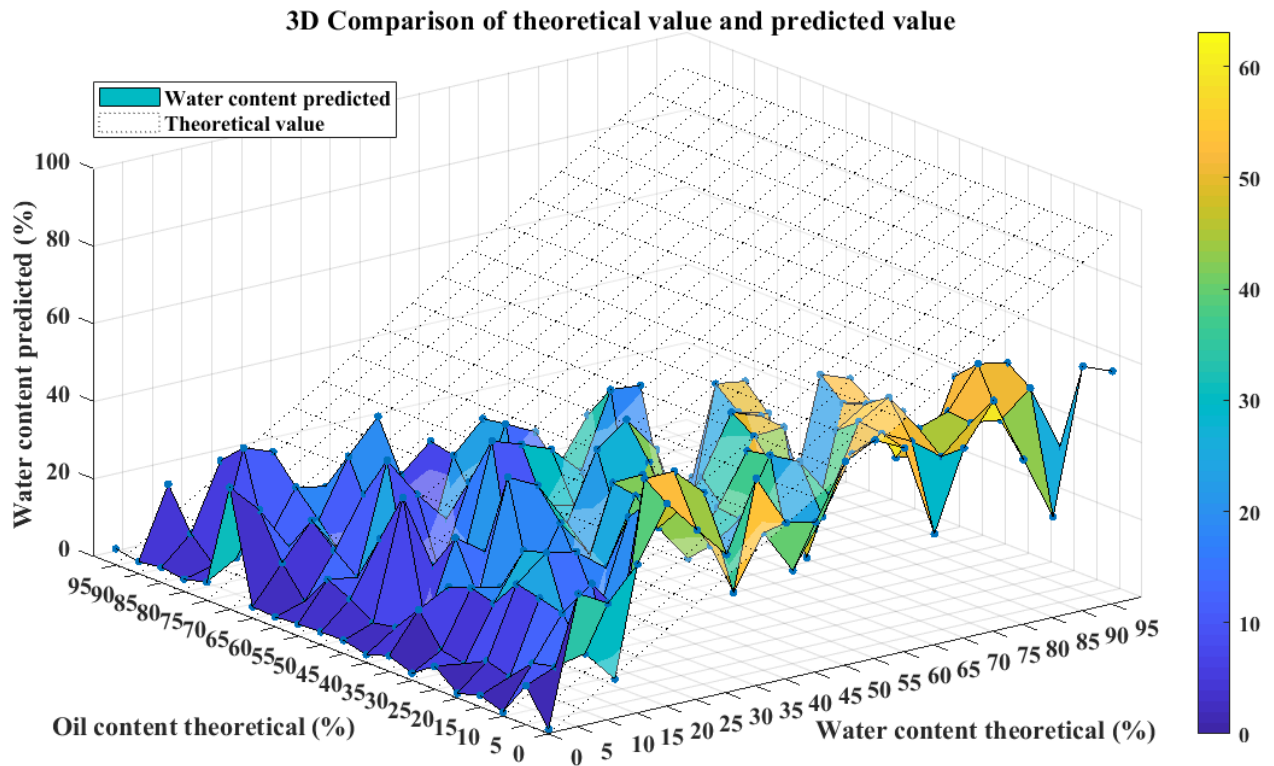The true content of oil in the pipeline, %

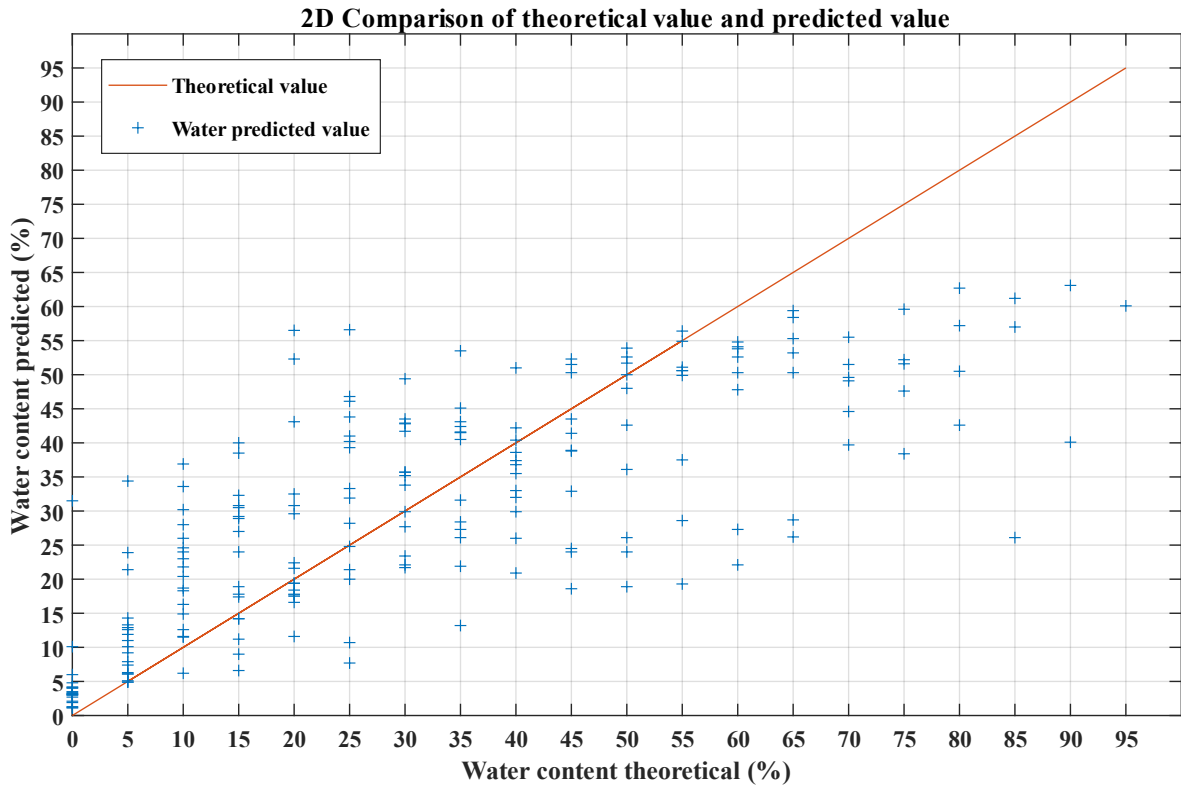Fig. 4.12 3D Comparison of theoretical value and predicted value (Table 4.11).

Fig. 4.13 2D Comparison of theoretical value and predicted value (Table 4.11).

From the figures shows above, the prediction of water content is terrible. This means it is also difficult to predict the oil content by training the model with only the data of the vertical testing. In Section 4.5 these results will be discussed carefully.

Table 4.12 shows the SSE value of the data in Table 4.11.

Table 4.12 The SSE values in each content of water based on Table 4.11.

Theoretical value, %	0	5	10	15	20	25	30	35	40	45
SSE	64.77	100.6	184.9	161	211.5	242.2	93.74	117	78.42	174.5
Theoretical value, %	50	55	60	65	70	75	80	85	90	95
SSE	249.1	262.2	363	478.2	494.8	679	772	1607	1607	1218

The SSE value of all the data is 246.61 in this part.

4.5. Discussions

The results of the model training are satisfactory. Although a certain degree of overfitting occurs (validation loss is higher than train loss), it is normal in regression machine learning.

In the pre-processing of the data, the integration of the measurement data of horizontal testing and vertical testing is selected, and the integrated new data is imported into the model for training in order to improve the accuracy by increasing the data dimension. However, in order to reflect the effect of this practice, the measurement data of the horizontal testing or the vertical testing is also processed separately into the same size data as explained in the aforementioned text, and the same model is imported for learning and detection. The prediction results of the neural network trained with these input data are represented by Error Sum of Square as the figure of merit, and the results are shown in the following table:

Table 4.13 SSE values for oil/water predictions from different neural networks.

	Horizontal + Vertical		Horizontal only		Vertical only	
Liquid	Oil	Water	Oil	Water	Oil	Water
Total SSE	0.8072	40.50	26.60	166.57	575.67	246.61

Smaller SSE values mean smaller measurement errors [61]. Therefore, when the

SSE is close to 0, the model selection and fitting are better, and the data prediction is more successful.

The results of Section 4.2 indicate that the network is available, and a high accuracy neural network could be trained using horizontal and vertical testing data as input. Although there are limits on the prediction of water content, the demand at the application level is mainly an accurate prediction of oil content. The results of Section 4.3 and 4.2 are similar, indicating that it is also possible to train the model to achieve the prediction of water/oil content through the data of horizontal testing, but the accuracy of this method is much lower. The results of Section 4.4 indicate that the measurement data obtained in vertical testing cannot be used alone to train the model.

There are two possible reasons why these three well-trained neural networks have limits or inaccuracies in predicting liquid content:

(1) The vector network analyzer has a dynamic range in the experiment. The system dynamic range of the VNA is defined as the difference between the measured power available on the test port and on the noise floor of the receiver. When the liquid content reaches a certain value, the signal to noise ratio is limited. Therefore, when the liquid content is increases, the measurement of the signal gradually reaches the limit of the dynamic range of the VNA, so the measured value is not accurate. The solution to this problem is to replace the settings in the experiment with a better frequency range, more suitable antennas, or another vector network analyzer with a larger dynamic range.

(2) The accuracy of the model is still not enough. Although it can be seen from Section 4.2 that the oil content can be predicted from 0% to 95%, it is impossible to predict the value when the water content reaches above 60%, as the data would be too close to those below 60%, making the model indistinguishable. The solution to this problem is to increase the depth of the network.

CHAPTER 5 CONCLUSIONS AND FUTURE WORK

5.1. Summary of Key Contributions

In this thesis, after comparing the advantages and disadvantages of many current methods for measuring three-phase flow, a self-built neural network model is proposed to accurately predict water/oil content through the collation and analysis of microwave measurement data, which offers a novel and valuable method to industries. Several antennas are used to measure signal changes caused by the multiphase flow inside the pipe. A neural network model is designed and trained to analyze these test data, which can achieve accurate prediction of oil content. The advantage of the proposed method is that it only needs to train the model in advance, and the computation is fast when predicting oil content in the unknown materials inside the pipe. Therefore, the cost is low, the operation is simple, and this system can be considered almost real-time detection due to the very small delay.

Besides, compared with other three-phase flow detection methods, this work can accurately measure the oil content under high water content and avoid using the imaging technology to image high-water-content multiphase flow which is a challenge to many tomographic methods. In this work, the model of the convolutional neural network can be trained in advance by using the database, and then real-time measurement can be achieved through the application of the trained model.

The main contributions of this thesis are as follows:

(1) The principles of antennas and convolutional neural networks are explained and analyzed to examine how they can be combined, which also the underpinning theories of the subsequent experimental equipment selection, experimental method design, and data extraction and processing.

(2) The experimental method is designed. Through the analysis of the data obtained, a program for batch extracting the required data and processing it into the required size is written using MATLAB. After data processing, a convolutional neural network model for regression analysis is designed and debugged by using Python.

(3) The oil content prediction results, which other methods couldn't measure accurately in full liquid range, are all good if both horizontal and vertical experimental data were used. The SSE value is 0.81 which means the model selection and fitting are good and the data prediction is successful.

5.2. Limitations

The main limitations of this study are listed below:

(1) Due to the nature of the microwave, the material of the pipe needs to be non-conductive.

(2) This test only measures static and stratified liquid environments. Models that have been trained may not be suitable for predicting complex liquid environments that are agitated and mixed.

(3) The frequency range and antennas used in this experiment are specially selected for this pipe. The selection of frequency range as well as antenna's factors should be reconsidered for different pipes.

(4) Each different pipeline requires a separate database collection and model training. Attention to the type of liquid in multiphase liquid should be paid.

5.3. Future Work

The future research and development of this achievement is advised to focus on the following aspects:

(1) Design a dedicated antenna especially with high gain because currently available antennas on the market are not suitable enough for measuring liquids in pipes. Furthermore, the use of an amplifier in circuit can be a good idea to increase the signal-to-noise ratio.

(2) Design a convolutional neural network model specially for $4 \times 1000 \times 2$ input data to discuss the causes of the poor performance if only horizontal (or vertical) data was used.

(3) Experiment by using VNA with 4 ports or more. This allows multiple pairs of antennas to be used simultaneously to measure data for more accurate.

(4) Detect the agitated liquid environment and expand the database and data types

of the training model in order to predict more complex liquid environments.

(5) Optimize the neural network model and increase the number of layers to make the model deeper in order to achieve higher resolution of data and more accurate prediction results.

References

[1] http://www.sensorcity.co.uk/wp-content/uploads/2017/11/Unilever-Challenge-Statements.pdf

[2] R. Thorn, Geir A. Johansen, and Bjørn T. Hjertaker. "Three-phase flow measurement in the petroleum industry." *Measurement Science and Technology* 24.1 (2012): 012003.

[3] G. Falcone, Geoffrey Hewitt, and Claudio Alimonti. *Multiphase flow metering: principles and applications*. Vol. 54. Elsevier, 2009.

[4] Hewitt, G. F., A. Hall, and L. Pan. "Three-Phase Gas-Liquid-Liquid flow." *A New Challenge, First International Symposiun on Two-phase Flow Modelling and Experimentation,* Italy19951 (1997).

[5] Y. LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *nature* 521.7553 (2015): 436.

[6] Ghosh-Dastidar, Samanwoy, and Hojjat Adeli. "Spiking neural networks." *International journal of neural systems* 19.04 (2009): 295-308.

[7] S. Azizi, Ebrahim Ahmadloo, and Mohamed M. Awad. "Prediction of void fraction for gas–liquid flow in horizontal, upward and downward inclined pipes using artificial neural network." *International Journal of Multiphase Flow* 87 (2016): 35-44.

[8] L. Scheers. "Multiphase flow metering: on its way from nursing to mature technology." *Proc. Hydrocarbon Production Accounting Workshop* (Moscow, December 2008). 2008.

[9] C. P. Nemarich."Time domain reflectometry liquid level sensors." *IEEE instrumentation & measurement magazine* 4.4 (2001): 40-44.

[10] P. C. Magnusson, *et al.: Transmission lines and wave propagation*. CRC press, 2017.

[11] J. R. Petrick. *System for sensing levels and electrical characteristics of fluent materials*. U.S. Patent No 3,474,337, 1969.

[12] A. Cataldo, *et al.* "Microwave TDR for real-time control of intravenous drip infusions." *IEEE Transactions on Instrumentation and Measurement* 61.7 (2012): 1866-1873.

[13] E. I. Trenkal and A. G. Loshchilov. "Method for improving measurement

accuracy of multilayer environment levels using time domain reflectometry." *2017 International Siberian Conference on Control and Communications (SIBCON)*. IEEE, 2017.

[14] M. J. Da Silva, S. Thiele, L. Abdulkareem, B. J. Azzopardi, and U. Hampel, "High-resolution gas-oil two-phase flow visualization with a capacitance wire-mesh sensor", *Flow Meas. Instrum.*, vol. 21, no. 3, pp. 191-197, Sep. 2010.

[15] R. Kipping, R. Brito, E. Scheicher, and U. Hampel, "Developments for the application of the Wire-Mesh Sensor in industries", *Int. J. Multiph. Flow, v*ol. 85, pp. 86-95, Oct. 2016.

[16] H. F. Velasco Pena and O. M. H. Rodriguez, "Applications of wire-mesh sensors in multiphase flows", *Flow Meas. Instrum.,* vol. 45, pp. 255-273, Oct. 2015.

[17] H-M. Prasser, Arnd Böttger, and Jochen Zschau. "A new electrode-mesh tomograph for gas–liquid flows." *Flow measurement and instrumentation* 9.2 (1998): 111-119.

[18] J. Szumowski, Tomasz Rymarczyk, and Przemysaw Adamkiewicz. "Wire-Mesh Sensor for Invasive Imaging of Vertical and Horizontal Flows of Liquids and Gases." *2019 Applications of Electromagnetics in Modern Engineering and Medicine (PTZE)*. IEEE, 2019.

[19] W. Liu, C. Tan, and F. Dong. "A wire-mesh sensor for air-water two-phase flow imaging." *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings.* IEEE, 2015.

[20] A. Sujiwa. "Wire-mesh sensor (WMS) application for water slope detection." *2017 International Seminar on Sensors, Instrumentation, Measurement and Metrology (ISSIMM)*. IEEE, 2017.

[21] M. J. Da Silva, E. Schleicher, and U. Hampel. "Capacitance wire-mesh sensor for fast measurement of phase fraction distributions." *Measurement Science and Technology* 18.7 (2007): 2245.

[22] Z. Wu and H. Wang. "Microwave Tomography for Industrial Process Imaging: Example Applications and Experimental Results." *IEEE Antennas and Propagation Magazine* 59.5 (2017): 61-71.

[23] Marco A. Rodriguez Frias and W. Yang. "Dual-modality four-wire electrical capacitance and resistance tomography." *2018 IEEE International Conference on*

Imaging Systems and Techniques (IST). IEEE, 2018.

[24] W. Yang. "Tomographic imaging based on capacitance measurement and industrial applications." *2007 IEEE International Workshop on Imaging Systems and Techniques*. IEEE, 2007.

[25] S. M. Huang, *et al.* "Design of sensor electronics for electrical capacitance tomography." *IEE Proceedings G (Circuits, Devices and Systems)* 139.1 (1992): 83-88.

[26] W. Yang. "Hardware design of electrical capacitance tomography systems." *Measurement Science and Technology* 7.3 (1996): 225.

[27] W. Yang. "Charge injection compensation for charge/discharge capacitance measuring circuits used in tomography systems." *Measurement Science and Technology* 7.7 (1996): 1073.

[28] Z. Guo. "New normalization method of imaging data for electrical capacitance tomography." *2011 International Conference on Mechatronic Science, Electric Engineering and Computer (MEC)*. IEEE, 2011.

[29] G. E. Fasching and Nelson S. Smith Jr. "A capacitive system for three-dimensional imaging of fluidized beds." *Review of scientific instruments* 62.9 (1991): 2243-2251.

[30] M. Wang. "Seeing a new dimension—The past decade's developments On electrical impedance tomography." *Progress in natural science* 15.S1 (2005): 1-13.

[31] B. Wang, *et al.* "A novel electrical resistance tomography system based on C4D technique." *2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings*. IEEE, 2012.

[32] F. Dong, *et al.* "Development of single drive electrode electrical resistance tomography system." *IEEE transactions on instrumentation and measurement* 55.4 (2006): 1208-1214.

[33] C. Qiu, B. S. Hoyle, and F. J. W. Podd, "Engineering and application of a dual-modality process tomography system", *Flow Measurement and Instrumentation*, vol. 18, pp. 247-254, 2007.

[34] J. Sun and W. Yang, "A dual-modality electrical tomography sensor for measurement of gas-oil-water stratified flows", *Measurement*, vol. 66, pp. 150-160, 2015.

[35] Z. Wu and H. Wang. "Microwave Tomography for Industrial Process Imaging: Example Applications and Experimental Results." *IEEE Antennas and Propagation Magazine* 59.5 (2017): 61-71.

[36] L. E. Larsen and John H. Jacobi. "Microwave scattering parameter imagery of an isolated canine kidney." *Medical physics* 6.5 (1979): 394-403.

[37] J. C. Bolomey, et al. "Microwave diffraction tomography for biomedical applications." *IEEE Transactions on Microwave Theory and Techniques* 30.11 (1982): 1998-2000.

[38] J. Richmond. "Scattering by a dielectric cylinder of arbitrary cross section shape." *IEEE Transactions on Antennas and Propagation* 13.3 (1965): 334-341.

[39] S. A. Caorsi, G. L. Gragnani, and M. Pastorino. "Two-dimensional microwave imaging by a numerical inverse scattering solution." *IEEE Transactions on Microwave Theory and Techniques* 38.8 (1990): 981-980.

[40] A. Franchois and C. Pichot. "Microwave imaging-complex permittivity reconstruction with a Levenberg-Marquardt method." *IEEE Transactions on Antennas and Propagation* 45.2 (1997): 203-215.

[41] A. Roger. "Newton-Kantorovitch algorithm applied to an electromagnetic inverse problem." *IEEE Transactions on Antennas and Propagation* 29.2 (1981): 232-238.

[42] A. Franchois and A. G. Tijhuis. "A quasi-Newton reconstruction algorithm for a complex microwave imaging scanner environment." *Radio Science* 38.2 (2003): 12-1.

[43] P. Lobel, *et al.* "Conjugate gradient method for solving inverse scattering with experimental data." (1996).

[44] J. Hu, *et al.* "Sequential quadratic programming method for solution of electromagnetic inverse problems." *IEEE transactions on antennas and propagation* 53.8 (2005): 2680-2687.

[45] S. J. Russell and P. Norvig. *Artificial intelligence: a modern approach.* Malaysia; Pearson Education Limited, 2016.

[46] J. Schmidhuber. "Deep learning in neural networks: An overview." *Neural networks* 61 (2015): 85-117.

[47] F. Rosenblatt. *The perceptron, a perceiving and recognizing automaton Project Para.* Cornell Aeronautical Laboratory, 1957.

[48] C. Farabet, *et al.* "Learning hierarchical features for scene labeling." *IEEE transactions on pattern analysis and machine intelligence* 35.8 (2012): 1915-1929.

[49] Y. LeCun, *et al.* "Handwritten digit recognition with a back-propagation network." *Advances in neural information processing systems.* 1990.

[50] Y. LeCun, *et al.* "Gradient-based learning applied to document recognition." *Proceedings of the IEEE*86.11 (1998): 2278-2324.

[51] D. H. Hubel and Torsten N. Wiesel. "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex." *The Journal of physiology* 160.1 (1962): 106-154.

[52] C. F. Cadieu, *et al.* "Deep neural networks rival the representation of primate IT cortex for core visual object recognition." *PLoS computational biology*10.12 (2014): e1003963.

[53] K. Fukushima and S. Miyake. "Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position." *Pattern recognition* 15.6 (1982): 455-469.

[54] L. Bottou, *et al.* "Experiments with time delay networks and dynamic time warping for speaker independent isolated digits recognition." *First European Conference on Speech Communication and Technology.* 1989.

[55] A. Waibel, *et al.* "Phoneme recognition using time-delay neural networks." *IEEE transactions on acoustics, speech, and signal processing* 37.3 (1989): 328-339.

[56] P. Y. Simard, David Steinkraus, and John C. Platt. "Best practices for convolutional neural networks applied to visual document analysis." *Icdar.* Vol. 3. No. 2003. 2003.

[57] R. Vaillant, Christophe Monrocq, and Y. LeCun. "Original approach for the localisation of objects in images." *IEE Proceedings-Vision, Image and Signal Processing* 141.4 (1994): 245-250.

[58] S. Lawrence, *et al.* "Face recognition: A convolutional neural-network approach." *IEEE transactions on neural networks* 8.1 (1997): 98-113.

[59] H. Yi and K. Boyle. *Antennas: from theory to practice.* John Wiley & Sons, 2008.

[60] R. E. Collin, *Antennas and radiowave propagation.* McGraw-Hill, 1985.

[61] F. Vogt. "A self‑guided search for good local minima of the sum‑of‑squared‑error in nonlinear least squares regression." *Journal of Chemometrics* 29.2 (2015): 71-79.