

## Article

# ViT-Based Image Regression Model for Shear-Strength Prediction of Transparent Soil

Ziyi Wang <sup>1</sup>, Jinqing Jia <sup>1,\*</sup> , Lihua Zhang <sup>1</sup> and Ziqi Li <sup>2</sup>

<sup>1</sup> School of Civil Engineering, Dalian University of Technology, Dalian 116024, China; ziyi.wang@mail.dlut.edu.cn (Z.W.); lihua2018@dlut.edu.cn (L.Z.)

<sup>2</sup> School of Civil Engineering, Nanjing Tech University, Nanjing 211816, China; liziqi@njtech.edu.cn

\* Correspondence: jiajq@dlut.edu.cn

**Abstract:** The direct-shear test is the primary method used to test the shear strength of transparent soil, but this experiment is complex and easily influenced by experimental conditions. In order to simplify the process of obtaining the shear strength of transparent soil, an image regression model based on a vision transformer (ViT) is proposed in this paper; this is used to recognize the shear strength of the soil based on images of transparent-soil patches. This model uses a convolutional neural network (CNN) to decompose the transparent-soil images into multiple image patches containing high-order features, utilizes a ViT for feature extraction, and designs a regression network to facilitate the transfer of information between the abstract image features and shear strength. This model solves the problem of boundary blurring and difficult-to-identify features in speckle images. To demonstrate the effectiveness of the proposed model, different parameters related to transparent soil were obtained by controlling the particle size of fused quartz sand and the content of aerosol; in addition, the friction angle and cohesive force of the transparent soil under different proportions were measured using direct-shear tests, serving as two datasets. The results show that the proposed method achieves correlations of 0.93 and 0.94 in the two prediction tasks, thus outperforming existing deep learning models.

**Keywords:** transparent soil; shear-strength prediction; deep learning; vision transformer; direct-shear tests



**Citation:** Wang, Z.; Jia, J.; Zhang, L.; Li, Z. ViT-Based Image Regression Model for Shear-Strength Prediction of Transparent Soil. *Buildings* **2024**, *14*, 959. <https://doi.org/10.3390/buildings14040959>

Academic Editor: Bjorn Birgisson

Received: 5 March 2024

Revised: 26 March 2024

Accepted: 29 March 2024

Published: 1 April 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Transparent soil is prepared by mixing aggregates with a pore fluid that has a similar or the same refractive index [1,2]. Conducting indoor model experiments using transparent-soil material [3–6] in combination with particle image velocimetry technology enables the deformation and evolution process within the soil to be observed [7–10]. Many scholars use transparent soil to study the stability of slopes [11,12], the shear deformation of soil [13], and the surface uplift in model experiments [14]. Preliminary studies indicate that transparent soil prepared using fused quartz sand as the soil framework [15], aerosol silica powder as the binder, and mixed mineral oil as the pore fluid exhibits mechanical characteristics that are similar to clay; it can therefore be used to create physically complex models [16]. To determine the shear strength of transparent soil, it is typically necessary to conduct direct-shear tests [17,18]. This process involves layering the prepared soil sample into a shear box according to the mass, compacting it, applying vertical pressure, removing the retaining pins, and finally conducting the shear test. This experiment often requires researchers to surrender a significant amount of their time, but it is a necessary step in experimental research. In order to obtain more accurate data, scholars both domestically and internationally have improved this experiment by designing a large-scale direct-shear apparatus and conducting corresponding experiments [19]. However, the direct-shear test method has certain limitations:

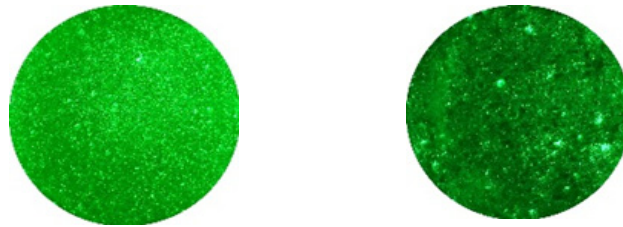
- Traditional direct-shear tests or triaxial direct-shear tests require multiple repetitions to be performed to obtain the average value of each set of test results [20–22], and the experimental procedures are relatively cumbersome. Additionally, the testing equipment often has certain requirements.
- In direct-shear tests, the specimen size is small; this enables the drainage conditions, stress conditions, compaction degree, shear rate, and other parameters of the soil sample to be controlled [23]. However, in large-scale geotechnical model experiments, it is difficult to ensure the uniformity of the soil at various locations within the same model. Furthermore, it is challenging to ensure that the moisture content and compaction degree of each group of soil specimens are the same in comparative tests. Therefore, there may be some discrepancies between the soil mechanics properties obtained under large-scale model experimental conditions and the data obtained under standard test conditions.

With the advancement of deep learning technology [24–26], its excellent performance in the field of image recognition has attracted widespread attention [27–29]. Muhammad et al. proposed a new approach for the prediction of  $D_h$  using gene expression programming [30]. This method can effectively extract features related to images and objects and perform tasks such as classification and regression. The shear strength of transparent soil may be connected to the distribution of optical spots in the images. The distribution of optical spots and the transparency differ between different transparent soils with different levels of shear strength. However, quantifying these differences is difficult, and traditional image processing methods are not effective for analyzing them. Therefore, using deep learning to analyze the optical speckle images of transparent soil and predict the shear strength of the soil in that area has strong applicative potential. The implementation of this research could enable human errors and experimental errors caused by the experimental conditions to be avoided, simplify the experimental process, and enhance the reliability of the measurement results.

The recognition of transparent soil relies heavily on deep learning feature extractors. The current mainstream feature extractors employed are convolutional neural networks (CNNs) [31–33] and vision transformers (ViTs) [34,35]. The core principle of a CNN is local perception, where the convolution operation allows the network to focus on local regions of the input image. By convolving the input image with convolutional kernels, CNNs extract features from the input image. This local perception enables the network to effectively capture the local structures and patterns of the image [36]. Compared to traditional CNN models, a ViT can better handle large-scale image data. The vision transformer is an image classification model that is based on the self-attention mechanism and was introduced by Dosovitskiy et al. [37] in 2020. Traditional CNNs have achieved great success in image classification tasks, but their convolutional operations can introduce some local biases when processing images; they also cannot maintain computational efficiency when dealing with large-sized images. A ViT treats the pixels in the image as a sequence and uses the self-attention mechanism to process this sequence, achieving comparable or even superior performance in image classification tasks compared to a CNN.

Finding features related to shear strength in the speckle images of transparent soil is a challenging task. It requires a model that not only has excellent local feature-capturing capabilities, but also understands the global features of the image. As shown in Figure 1, unlike existing image classification tasks, the features in the speckle image of transparent soil are distributed globally rather than concentrated in specific areas. Additionally, the speckles do not have significant boundaries, making it difficult for existing methods to capture their features. Based on previous research, CNN-based feature-processing modules are able to perceive details well, but remain limited in their ability to capture global features [38]. A ViT has excellent global feature-capturing capabilities [39]; however, its image preprocessing method, which involves the segmentation of image blocks, does not convert the detailed information of the image into high-level features and is not suitable for processing transparent-soil images. Furthermore, predicting the shear strength of trans-

parent soil based on images is, essentially, a regression task. In image processing problems, deep learning does not perform as well in regression tasks as it does in classification tasks. Therefore, it is necessary to develop a regression module that is able to facilitate the transfer of information between the feature extraction of transparent soil and shear-strength prediction processes.

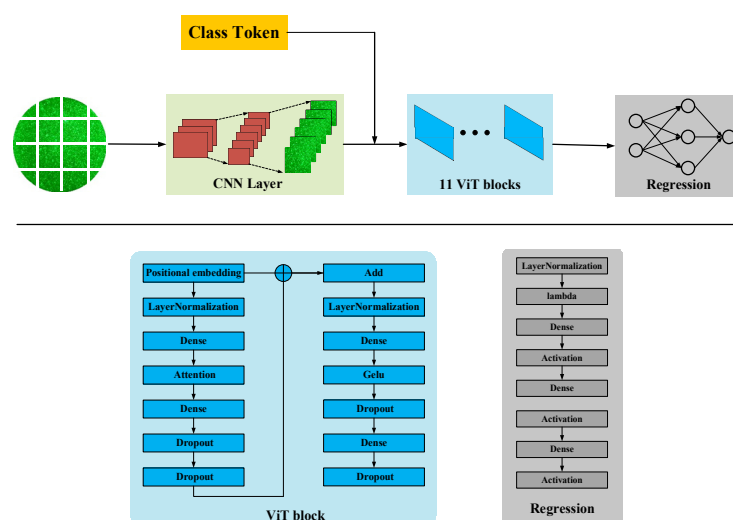


**Figure 1.** Transparent-soil speckle images with different parameters.

In this paper, a novel image regression model called the ViT-based image regression model (VIRM) is proposed; this model aims to improve the poor performance of existing methods in transparent-soil image feature extraction tasks. The input images are pre-processed using a CNN module, and the segmented image patches are replaced with feature maps to enter the transformer encoder. After the transformer, they are concatenated with the regression module to predict the shear strength. To demonstrate the effectiveness of the proposed model, transparent soils with different parameters are obtained by controlling the particle size of fused quartz sand and the content of fumed silica powder. The shear strength and cohesion of the transparent cemented soils under different ratios are measured using direct-shear tests. The soil samples under the same ratio are captured using laser light to create two transparent-soil datasets, which prove the effectiveness of the proposed method. Additionally, the proposed method does not require the addition of tracer particles, but only uses the light spots reflected by the fused quartz sand as tracer particles.

## 2. Methodology

As shown in Figure 2, the proposed VIRM comprises a CNN, ViT blocks and a regression module. The CNN and ViT blocks form the component used to perform the image feature extraction. The regression module primarily consists of fully connected layers and activation functions. It is used to unfold the abstract features of the image and predict the mechanical performance via regression. Each of these modules will be described in detail later.



**Figure 2.** Flow chart of the proposed method.

### 2.1. Image Feature Extraction

The structure of the component used for image feature extraction is shown in Table 1. The scattered image of the transparent soil first travels through a CNN layer, whose main purpose is to convert the scattered image into a multi-dimensional feature map that contains the multi-level features of the scattered image. The feature map is then passed into the ViT module. The conventional ViT module splits the complete image into  $N$  blocks of the same size, and then the  $N$  blocks are converted into  $N$  high-dimensional image block feature vectors via a linear mapping layer; however, such methods are limited in their ability to deal with the positional relationships between image blocks [40]. Since the CNN has the property of inductive bias and the transformer has the ability to perform strong global inductive modeling, better results can be obtained when using the hybrid CNN + transformer model; therefore, instead of image segmentation, the CNN is used in this paper.

**Table 1.** The structure of the feature extraction component.

Layer Name	Input Size	Output Size
CNN	(224, 224, 3)	(197, 768)
LayerNormalization	(197, 768)	(197, 768)
Dense	(197, 768)	(197, 2304)
Attention	(197, 2304)	(197, 768)
Dense	(197, 768)	(197, 768)
Dropout	(197, 768)	(197, 768)
Dropout	(197, 768)	(197, 768)
Add	(197, 768), (197, 768)	(197, 768)
LayerNormalization	(197, 768)	(197, 768)
Dense	(197, 768)	(197, 3072)
Gelu	(197, 3072)	(197, 3072)
Dropout	(197, 3072)	(197, 3072)
Dense	(197, 3072)	(197, 768)
Dropout	(197, 768)	(197, 768)

A ViT is an image classification model based on the transformer model. Its principles are as follows: (a) Patch embedding—the input image is divided into uniform patches, and each patch is flattened into a low-dimensional vector. These patch-embedding vectors are then passed to the transformer model as input. (b) Position embedding—a position encoder is used to generate position-embedding vectors for each patch. Position-embedding vectors provide the input patches with global positional information. (c) Encoder—the patch embeddings and position embeddings are passed to the transformer encoder as input. (d) Decoder—the output of the transformer encoder is passed to the decoder, which comprises a fully connected layer and a SoftMax layer. The ViT model completes its image classification and object detection tasks through the self-attention mechanism and feed-forward neural networks of the transformer.

The core component of a ViT is the attention mechanism, which forces the model to focus on more important feature maps [41,42]. Because different feature maps contribute differently to the prediction task, weights are added to each feature map as an indication of the importance of the feature map. The attention mechanism used in this research is self-attentive, where a weight  $w_{attention}$  of the same dimension is assigned to the input vector  $X$ . Through multiple rounds of training, this weight can represent the values that are more important to the convergence of the model with the input vector. This is achieved by performing a similarity calculation using the dot product, as follows:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

where  $Q$ ,  $K$ , and  $V$  denote “query”, “key”, and “value”, respectively;  $d_k$  is the scaling factor and denotes the dimensionality of  $K$ . For larger values of  $d_k$ , the dot product is too large, thus pushing the SoftMax function to regions with very small gradients. To counteract this effect, the dot product is scaled using  $\frac{1}{\sqrt{d_k}}$ .

## 2.2. Regression Module

As shown in Table 2, the regression module is a linear layer that connects the feature extraction component to the labels, and mainly consists of activation functions and dense layers. The feature extraction component converts the image features into feature vectors, which are first normalized and then passed through several activation and dense layers, which increase the nonlinearity of the regression module; finally, the image features are connected to the labels.

**Table 2.** Structure of the regression module.

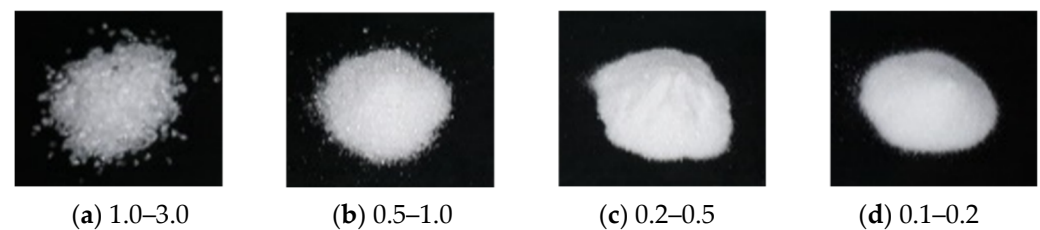
Layer Name	Input Size	Output Size
LayerNormalization	(197, 768)	(197, 768)
Lambda	(197, 768)	(None, 768)
Dense	(None, 768)	(None, 128)
Activation	(None, 128)	(None, 128)
Dense	(None, 128)	(None, 256)
Activation	(None, 256)	(None, 256)
Dense	(None, 256)	(None, 1)
Activation	(None, 1)	(None, 1)

## 3. Transparent-Soil Straight-Shear Experiment

### 3.1. Transparent Cemented Soil Preparation

The transparent cemented soil specimens were formulated from fused silica, a refractive-index-matched pore solution, and nanoscale hydrophobic fumed silica powder. The hydrophobic fumed silica powder was used as a binder with refractive-index-matched fused quartz sand (1.4585); this had the appearance of a white powder when dried. Four particle sizes of fused quartz sand were selected as the soil skeleton for the test, with a particle density of 2300 kg/m<sup>3</sup> and particle sizes of 0.1–0.2 mm, 0.2–0.5 mm, 0.5–1.0 mm, and 1.0–3.0 mm. The preparation ratios of different proportions of pore fluid were compared and the refractive index of the mixed solution was measured using an Abbe refractometer. It was found that at 25 °C, when n-dodecane and 15# white mixed mineral oil were mixed at a mass ratio of 1:20, the refractive index of the pore fluid reached 1.4585; this resulted in soil with the best transparency.

The preparation of the transparent-soil specimen was conducted as follows: (a) The specimens were washed with water to remove impurities, and then put into a drying oven for drying. (b) The pore solution was prepared by mixing n-dodecane and 15# white oil in proportion so that the refractive index of the mixture was 1.4585. (c) The fused quartz sand, silica powder and pore solution were weighted in proportion to each other. The quartz sand and silica powder were mixed and stirred well so that the silica powder was adsorbed onto the surface of the quartz sand particles. Finally, the weighed pore liquid was added to the mixture and stirred well. Because air was mixed into the cemented transparent soil at this time, it was milky white or translucent. (d) The transparent cemented soil made in the previous step was compacted in layers into a test tube that was 150 mm long and 25 mm in diameter. The test tube was then placed in a vacuum chamber and evacuated for 30 min. After extracting the gas inside the soil, the soil particles were rearranged under the action of atmospheric pressure, and the compactness reached 70%. The configured specimen is shown in Figure 3.



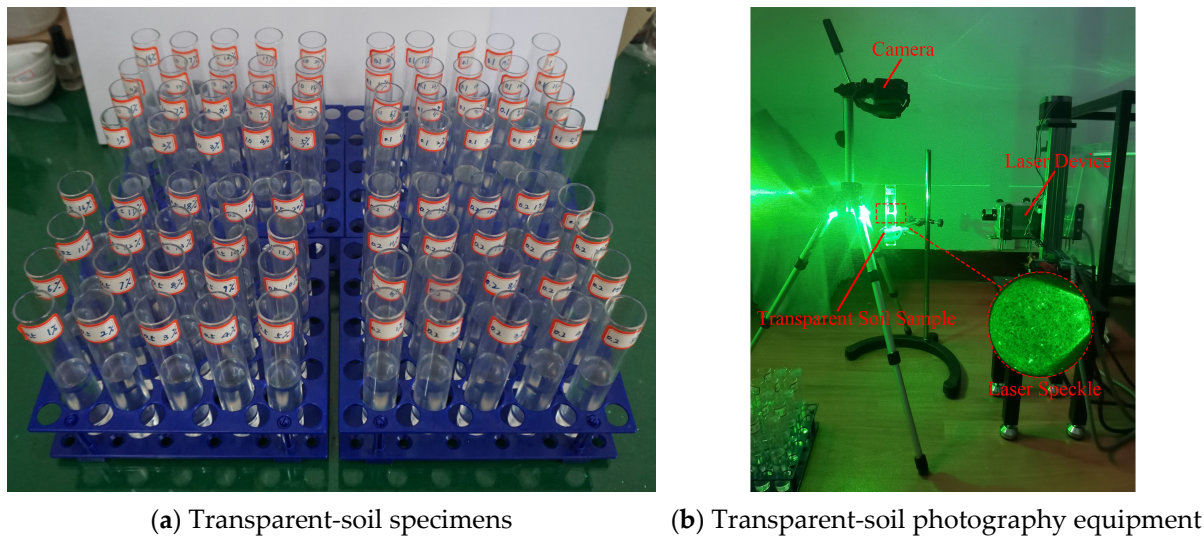
**Figure 3.** Fused quartz sand.

### 3.2. Scattered Image Acquisition

As shown in Table 3, to obtain transparent cemented soil patches comprising fused quartz sand with different particle sizes and silica powder contents, four different quartz sands with grain sizes of 0.1–0.2 (fine sand), 0.2–0.5 (medium sand), 0.5–1.0 (coarse sand), and 1.0–3.0 (fine gravel) were selected; the fumed silica powder was added to each quartz sand grain size, and the content of fumed silica powder was increased from 0% to 20%. As shown in Figure 4, a scatter collection test was performed on the transparent cemented soil; this was conducted in a dark room using a 2000 mW and 532 nm sheet laser to irradiate the transparent cemented soil. After several adjustments and comparisons, the soil speckle field brightness was moderate, with a 4.0 w laser intensity, and the speckle field distribution was uniform. The test was therefore suitable for observing and acquiring the speckle field image.

**Table 3.** Speckle image of transparent soil with different parameters.

Content of Fumed Silica Powder	Particle Size of Fused Quartz Sand/mm			
	0.1–0.2	0.2–0.5	0.5–1.0	1.0–3.0
0%				
5%				
10%				
15%				
20%				

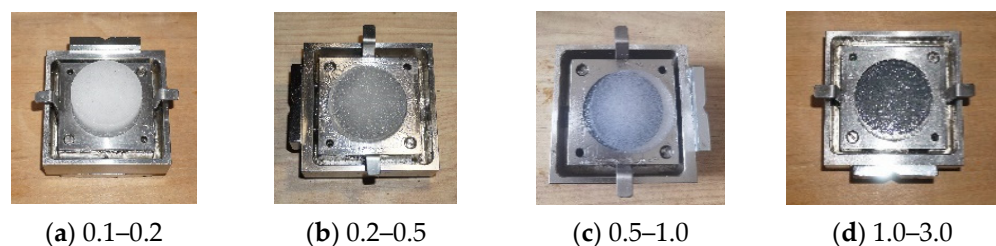


**Figure 4.** Transparent-soil specimens and transparent-soil photograph equipment.

### 3.3. Transparent-Soil Straight-Shear Tests

To determine the shear strength of the transparent cemented soils with different mix ratios, straight-shear tests were conducted on all samples. The experimental setup utilized a conventional strain-controlled direct-shear apparatus. The soil samples with a diameter of 61.8 mm and an initial height of 20 mm were compacted to the corresponding compactness. The consolidated quick shear test method was then performed, and a shear rate of 0.8 mm/min was used.

When the percentage of hydrophobic fumed silica powder is  $<3\%$ , the viscosity of the soil is extremely low, and it is difficult to create a shape. Therefore, when the percentage of fumed silica powder was  $<3\%$ , the method used to create samples of sand-like soil was employed; subsequently, the sample was filled into the shear box in layers according to the mass and then compacted. When the percentage of hydrophobic fumed silica powder was  $>3\%$ , the method used to create samples of clay-like soil was employed for the shear test. The configured transparent cemented soil was then put into the mold in layers according to the calculated mass and compacted; subsequently, the specimen was pushed into the shear box after being cut with the ring knife. Because the specimen was not fully saturated, the specimen was translucent or light white; the sheared specimen is shown in Figure 5. Although the transparent cemented soil used in the model box test was fully saturated, the transparent-soil models were all small in size due to the visible depth of the soil. Therefore, the pore pressure in the models was almost negligible, and unsaturated soil specimens with a higher saturation could be used as approximate substitutes for saturated soil specimens in this test. Finally, the shear strength of the soil was calculated from the following measured parameters: the cohesion and friction angle.



**Figure 5.** The sheared transparent-soil specimen.

#### 3.3.1. Dataset 1: Cohesion

Cohesion refers to the force generated by the cementation and electrostatic gravitational force between particles; this primarily depends on the physicochemical interaction

between soil particles and specifically the intersection of the cementing substances between the particles. In this experiment, both the soil density and quartz sand particle size were the same; therefore, the percentage of hydrophobic fumed silica powder was the main factor affecting the cohesion of the specimen. In this dataset, the proposed ViT model was trained using the captured transparent-soil images as the input images and the cohesion as the label; the dataset contained a total of 2000 pre-processed samples.

### 3.3.2. Dataset 2: Friction Angle

The friction angle is an important index used to describe the soil friction strength. It is influenced by two main factors, namely the sliding friction between particles and the interlocking friction between particles. In this experiment, the influence of the percentage of hydrophobic fumed silica powder on the friction angle of the specimen was small, as the silica powder was in the form of gel after mixing with the pore liquid. The higher the content of quartz sand, the less pore space there is between the particles; this leads to the enhanced self-locking effect and occlusion between the particles, resulting in a larger friction angle. Therefore, in this paper, the content of quartz sand was the main factor affecting the friction angle of transparent cemented soil. Secondly, quartz sand particle size has a significant impact on the friction angle, as it affects the magnitude of interparticle occlusion. In this dataset, 2000 transparent-soil cross-section images were used as the input of the dataset, and the corresponding friction angles were used as labels to train the proposed ViT.

## 4. Experiments and Results

### 4.1. Experiment Configuration

The network model used in this paper was trained on a high-performance GPU NVIDIA GeForce RTX2080. The software environment was CUDA version bit 10.2, the python version was 3.6.3, the system version was WIN10, and the network used Tensorflow version 1.13. During the training period, the input size of the images was  $224 \times 224$ . In this study, the network was trained using migration learning and freezing in order to improve the efficiency and accuracy of the training. In this study, the network was trained using a total of 500 epochs; the first 200 epochs were trained frozen and the last 300 epochs were trained unfrozen until the network converged. In the frozen phase, the batch was set to 32; in the unfrozen phase, the batch was set to 16. The initial learning rate of the network was set to  $1 \times 10^{-2}$ , and the minimum learning rate of the network was 0.01 times the initial learning rate. The network parameters were all updated according to the adaptive moment estimation (Adam) optimization method.

### 4.2. Experimental Results

In this section, some state-of-the-art backbone networks are used for comparison to demonstrate the excellent performance of the proposed feature extraction model (backbone network part) on this dataset. We used the prediction error as a metric, specifically the difference between the predicted and true values. The statistical results are plotted as histograms and a normal fit curve is attached.

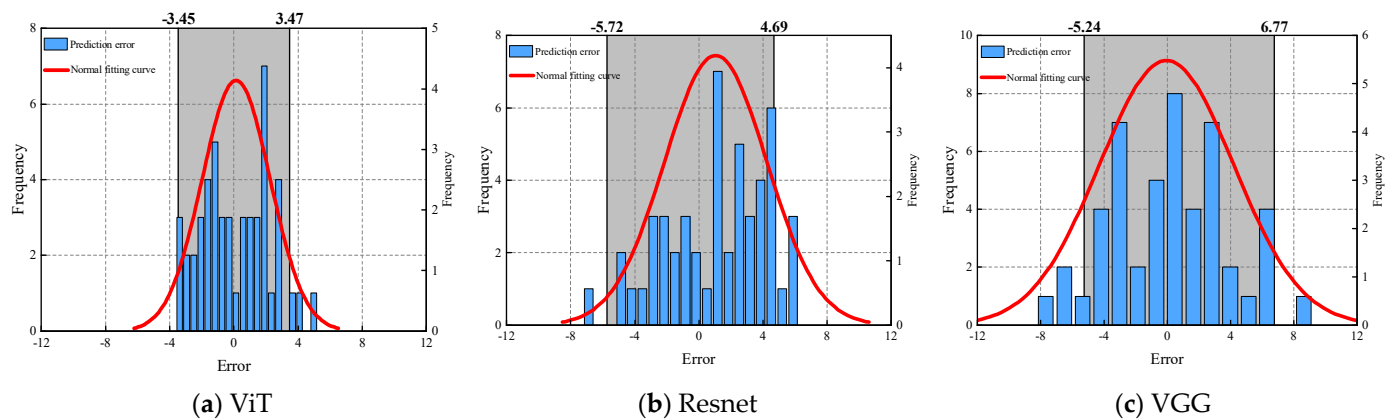
The backbone networks that were compared are VGG and ResNet, both of which have powerful feature extraction capabilities and have successfully achieved image classification and regression in many scenarios. The VGG network is a deep convolutional neural network model that was proposed by researchers from the University of Oxford. It is characterized by the use of very small convolutional kernels (typically  $3 \times 3$ ) and a very deep network structure. This design enables the VGG network to use a smaller number of parameters. ResNet addresses the gradient problem by introducing shortcut connections that span across network layers. Shortcut connections directly pass the input information (i.e., intermediate features) around various layers to the subsequent layers, allowing the neural network to better optimize residual information. This type of shortcut connection enables ResNet to easily train deep networks without experiencing performance

degradation. The training times and average prediction errors for different methods are shown in Table 4.

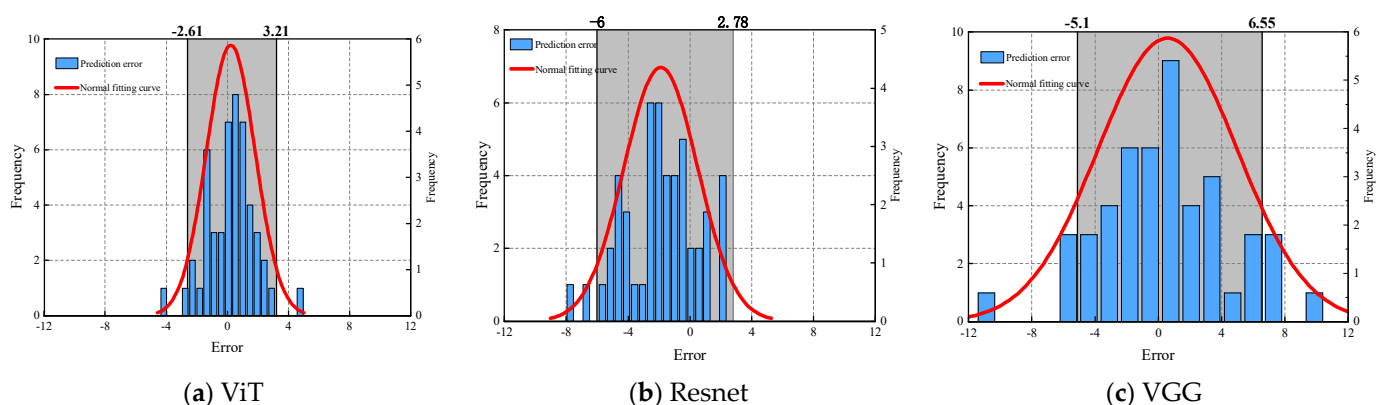
**Table 4.** Training times and average prediction errors for different methods.

Method	Training Time	Cohesion Average Error	Friction Angle Average Error
Vit	124 min	1.64	0.73
Resnet	78 min	3.95	2.33
VGG	55 min	5.27	2.26

As shown in Figure 6, the prediction errors of the three backbone networks were recorded in the cohesion prediction task. It can be observed from the distribution curve of the fitted errors and the 95% confidence interval that the prediction accuracy of the ViT is higher than that of the other two backbone networks, with a 95% confidence interval of  $(-3.45, 3.47)$ ; this is significantly narrower than the confidence intervals of Resnet and VGG, which are  $(-5.72, 4.69)$  and  $(-5.24, 6.77)$ , respectively. The prediction results of the friction angle, as shown in Figure 7, indicate that the predictive accuracy of the ViT is slightly better than that of Resnet and significantly higher than that of VGG.



**Figure 6.** Error statistics for cohesion prediction tasks with different backbone networks.



**Figure 7.** Error statistics for friction angle prediction tasks with different backbone networks.

In total, 50 samples were randomly selected for testing, and the proposed method was used for prediction. Resnet and VGG were used as control groups. The test results, as shown in Figures 8 and 9, indicate that the results predicted by the proposed ViT are highly consistent with the ground truth. Compared to Resnet and VGG, the error was significantly reduced. Using the Spearman correlation coefficient to assess the correlation between the ViT and the ground truth, the results show that the correlation between the ViT's

predictions reached 0.93 and 0.94 in both prediction tasks; this indicates that the reliability of the proposed method is strong. By analyzing the normalized covariance matrix of the two prediction targets, we can observe that the proposed ViT achieves a correlation of 0.99 with the ground truth, surpassing the reliability of the other two methods. The results are shown in Figure 10. Therefore, this method demonstrates significant applicative potential in predicting the mechanical properties of soil.

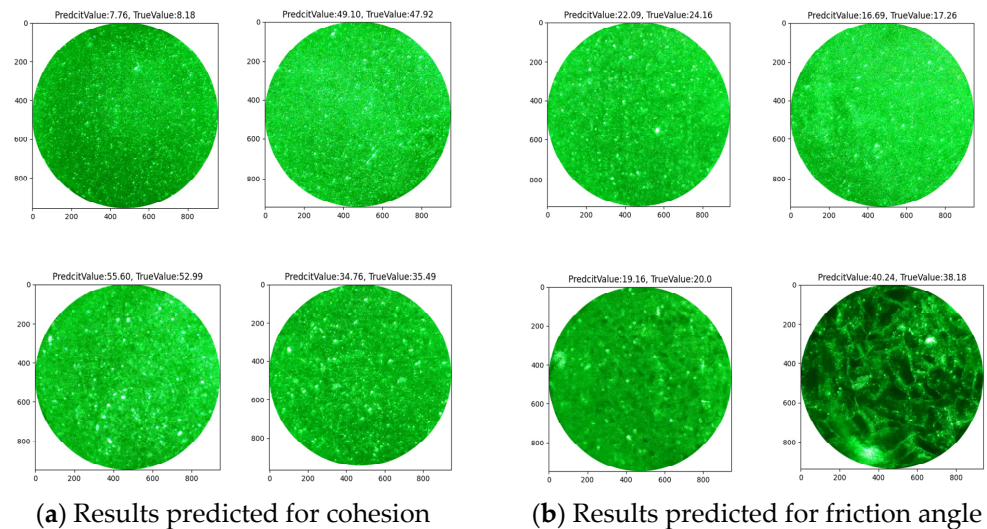


Figure 8. Results predicted for cohesion and friction angle.

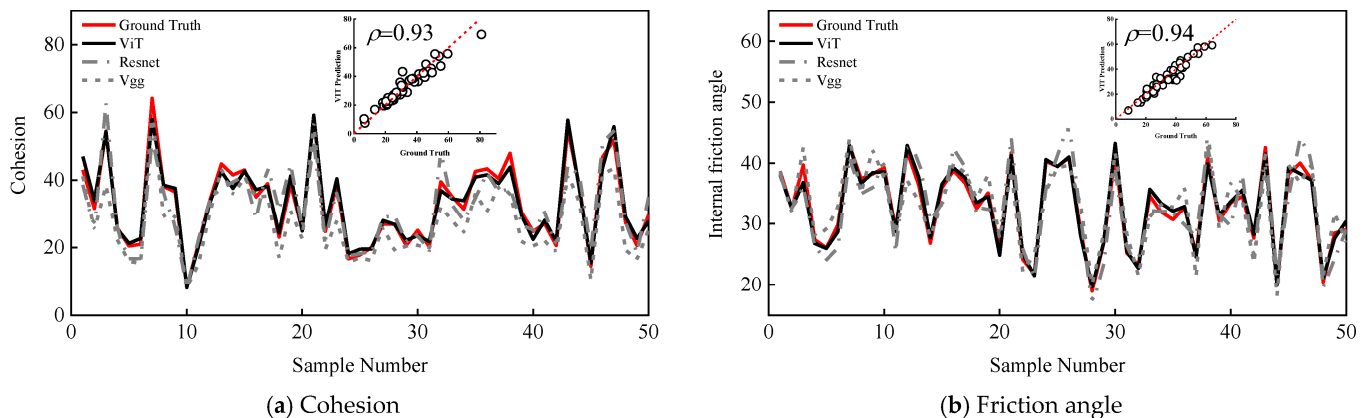


Figure 9. Prediction results for 50 samples and correlation statistics.

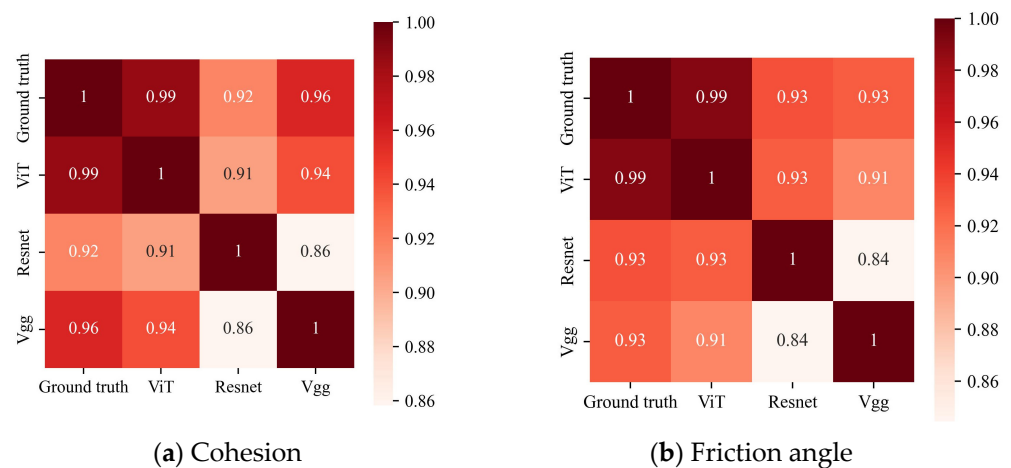


Figure 10. The correlation matrix of different methods.

## 5. Conclusions

In order to predict the shear strength of transparent soil based on images, a VIRM image regression prediction model is proposed in this paper. This model mainly comprises a CNN image preprocessing module, a ViT module, and a regression module. To address the issue of unclear feature boundaries in transparent-soil images and the characteristics of their global distribution, a combination of CNN and ViT is proposed; this overcomes the limitations of traditional ViT models, which rely on the input of segmented soil blocks. A regression module is also proposed to facilitate the transfer of information between transparent-soil feature extraction and shear-strength prediction processes.

To demonstrate the effectiveness of the proposed method, transparent-soil samples with different parameters were obtained by controlling the particle size of fused quartz sand and the content of fumed silica. The shear strength and cohesion of transparent cemented soil with different proportions were measured via direct-shear tests. Laser spot acquisition was performed on the soil samples under controlled conditions, resulting in the creation of two transparent-soil datasets. Via the performance of tests, the following conclusions were drawn:

- (1) To demonstrate the effectiveness of the proposed feature extraction module, the module was replaced with classical Resnet and VGG models for comparison. The results showed that in both datasets, the feature extraction module based on CNN + ViT was more suitable for predicting the shear strength of transparent soil. The shear-strength prediction method based on this module achieved a smaller error distribution and higher prediction accuracy.
- (2) Fifty samples were randomly selected for prediction, and the correlation coefficients between the predicted values and the true values were calculated. The results showed that the proposed method achieved correlation coefficients of 0.93 and 0.94 in the two datasets, indicating a high level of reliability.

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

**Funding:** The National Natural Science Foundation of China with Grant No. 52278332.

**Data Availability Statement:** The data presented in this study are available upon request from the corresponding author. The data are not publicly available due to privacy restrictions.

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

## References

1. Zhou, C.; Ma, W.; Sui, W. Transparent soil model test of a landslide with umbrella-shaped anchors and different slope angles in response to rapid drawdown. *Eng. Geol.* **2022**, *307*. [[CrossRef](#)]
2. Ezzein, F.M.; Bathurst, R.J. A new approach to evaluate soil-geosynthetic interaction using a novel pullout test apparatus and transparent granular soil. *Geotext. Geomembr.* **2014**, *42*, 246–255. [[CrossRef](#)]
3. Wang, Z.; Luo, G.; Kong, G.; Zhang, Y.; Lu, J.; Chen, Y.; Yang, Q. Centrifuge model tests on anchor pile of single point mooring system under oblique pullout load using transparent sand. *Ocean Eng.* **2022**, *264*, 112441. [[CrossRef](#)]
4. Liu, C.; Tang, X.; Wei, H.; Wang, P.; Zhao, H. Model Tests of Jacked-Pile Penetration into Sand Using Transparent Soil and Incremental Particle Image Velocimetry. *KSCE J. Civ. Eng.* **2020**, *24*, 1128–1145. [[CrossRef](#)]
5. Yuan, B.; Li, Z.; Zhao, Z.; Ni, H.; Su, Z.; Li, Z. Experimental study of displacement field of layered soils surrounding laterally loaded pile based on transparent soil. *J. Soils Sediments* **2021**, *21*, 3072–3083. [[CrossRef](#)]
6. Ma, S.; Duan, Z.; Huang, Z.; Liu, Y.; Shao, Y. Study on the stability of shield tunnel face in clay and clay-gravel stratum through large-scale physical model tests with transparent soil. *Tunn. Undergr. Space Technol.* **2021**, *119*, 104199. [[CrossRef](#)]
7. Chen, J.-F.; Gu, Z.-A.; Rajesh, S.; Yu, S.-B. Pullout Behavior of Triaxial Geogrid Embedded in a Transparent Soil. *Int. J. Géoméch.* **2021**, *21*, 04021003. [[CrossRef](#)]
8. Tao, F.-J.; Xu, Y.; Zhang, Z.; Ye, G.-B.; Han, J.; Cheng, B.-N.; Liu, L.; Yang, T.-L. Progressive development of soil arching based on multiple-trapdoor tests. *Acta Geotech.* **2022**, *18*, 3061–3076. [[CrossRef](#)]
9. Chen, Q.; Dong, G.C.; Wang, C.; Zhu, B.L.; Zhao, X.Y. Characteristics Analysis of Soil Arching Effect Behind Pile Based on Transparent Soil Technology. *J. Southwest Jiaotong Univ.* **2020**, *55*, 509–522.

10. Yuan, B.; Sun, M.; Xiong, L.; Luo, Q.; Pradhan, S.P.; Li, H. Investigation of 3D deformation of transparent soil around a laterally loaded pile based on a hydraulic gradient model test. *J. Build. Eng.* **2019**, *28*, 101024. [\[CrossRef\]](#)
11. Sui, W.; Zheng, G. An experimental investigation on slope stability under drawdown conditions using transparent soils. *Bull. Eng. Geol. Environ.* **2017**, *77*, 977–985. [\[CrossRef\]](#)
12. Lanting, W.; Qiang, X.; Shanyong, W.; Cuilin, W.; Xu, J. The morphology evolution of the shear band in slope: Insights from physical modelling using transparent soil. *Bull. Eng. Geol. Environ.* **2020**, *79*, 1849–1860. [\[CrossRef\]](#)
13. Li, Y.; Zhou, H.; Liu, H.; Ding, X.; Zhang, W. Geotechnical properties of 3D-printed transparent granular soil. *Acta Geotech.* **2020**, *16*, 1789–1800. [\[CrossRef\]](#)
14. Gong, Q.; Zhao, Y.; Zhou, J.; Zhou, S. Uplift resistance and progressive failure mechanisms of metro shield tunnel in soft clay. *Tunn. Undergr. Space Technol.* **2018**, *82*, 222–234. [\[CrossRef\]](#)
15. Zhong, W.; Liu, H.; Gu, D.; Zhang, W.; Yang, C.; Gao, X. Development of a preparation method of transparent soil-rock mixture for geotechnical laboratory modeling. *Eng. Geol.* **2022**, *301*, 106622. [\[CrossRef\]](#)
16. Leng, X.L.; Wang, C.; Pang, R.; Sheng, Q. Experimental study on the strength characteristics of a transparent cemented soil. *Rock Soil Mech.* **2021**, *42*, 2059.
17. Wei, L.; Xu, Q.; Wang, S.; Wang, C.; Chen, J. Development of transparent cemented soil for geotechnical laboratory modelling. *Eng. Geol.* **2019**, *262*, 105354. [\[CrossRef\]](#)
18. Yang, X.; Jin, G.; Huang, M.; Tang, K. Material preparation and mechanical properties of transparent soil and soft rock for model tests. *Arab. J. Geosci.* **2020**, *13*, 1–10. [\[CrossRef\]](#)
19. Vangla, P.; Gali, M.L. Effect of particle size of sand and surface asperities of reinforcement on their interface shear behaviour. *Geotext. Geomembr.* **2016**, *44*, 254–268. [\[CrossRef\]](#)
20. Peerun, M.I.; Ong, D.E.L.; Choo, C.S. Interpretation of Geomaterial Behavior during Shearing Aided by PIV Technology. *J. Mater. Civ. Eng.* **2019**, *31*. [\[CrossRef\]](#)
21. Zhang, G.; Liang, D.; Zhang, J.-M. Image analysis measurement of soil particle movement during a soil–structure interface test. *Comput. Geotech.* **2006**, *33*, 248–259. [\[CrossRef\]](#)
22. Shi, H.-L.; Zhao, L.-Y.; Zhu, Q.-Z. Experimental and analytical investigations on strength and deformation behaviour of red sandstone under conventional triaxial compression. *Eur. J. Environ. Civ. Eng.* **2023**, *27*, 1–15. [\[CrossRef\]](#)
23. Nam, S.; Gutierrez, M.; Diplas, P.; Petrie, J. Determination of the shear strength of unsaturated soils using the multistage direct shear test. *Eng. Geol.* **2011**, *122*, 272–280. [\[CrossRef\]](#)
24. Amiri, S.; Rajabi, A.; Shabanlou, S.; Yosefvand, F.; Izadbakhsh, M.A. Prediction of groundwater level variations using deep learning methods and GMS numerical model. *Earth Sci. Inform.* **2023**, *16*, 3227–3241. [\[CrossRef\]](#)
25. Fukuoka, T.; Fujiu, M. Detection of Bridge Damages by Image Processing Using the Deep Learning Transformer Model. *Buildings* **2023**, *13*, 788. [\[CrossRef\]](#)
26. Ali, L.; Al Jassmi, H.; Khan, W.; Alnajjar, F. Crack45K: Integration of Vision Transformer with Tubularity Flow Field (TuFF) and Sliding-Window Approach for Crack-Segmentation in Pavement Structures. *Buildings* **2022**, *13*, 55. [\[CrossRef\]](#)
27. Miao, S.; Wang, Z.J.; Liao, R. A CNN Regression Approach for Real-Time 2D/3D Registration. *IEEE Trans. Med. Imaging* **2016**, *35*, 1352–1363. [\[CrossRef\]](#)
28. Niu, Z.; Zhou, M.; Wang, L.; Gao, X.; Hua, G. Ordinal regression with multiple output cnn for age estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
29. Niyogi, A.; Ansari, T.A.; Sathapathy, S.K.; Sarkar, K.; Singh, T.N. Machine learning algorithm for the shear strength prediction of basalt-driven lateritic soil. *Earth Sci. Inform.* **2023**, *16*, 899–917. [\[CrossRef\]](#)
30. Raja, M.N.A.; Abdoun, T.; El-Sekelly, W. Smart prediction of liquefaction-induced lateral spreading. *J. Rock Mech. Geotech. Eng.* **2023**. [\[CrossRef\]](#)
31. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [\[CrossRef\]](#)
32. Guérin, J.; Thiery, S.; Nyiri, E.; Gibaru, O.; Boots, B. Combining pretrained CNN feature extractors to enhance clustering of complex natural images. *Neurocomputing* **2020**, *423*, 551–571. [\[CrossRef\]](#)
33. Varshni, D.; Thakral, K.; Agarwal, L.; Nijhawani, R.; Mittal, A. Pneumonia detection using CNN based feature extraction. In Proceedings of the 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), Coimbatore, India, 20–22 February 2019.
34. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021.
35. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 6000–6010.
36. Liu, Y.; Pu, H.; Sun, D.-W. Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices. *Trends Food Sci. Technol.* **2021**, *113*, 193–204. [\[CrossRef\]](#)
37. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv* **2020**, arXiv:2010.11929.

38. Song, L.; Xia, M.; Weng, L.; Lin, H.; Qian, M.; Chen, B. Axial Cross Attention Meets CNN: Bibranch Fusion Network for Change Detection. *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2022**, *16*, 21–32. [[CrossRef](#)]
39. Zhang, Z.; Lu, X.; Cao, G.; Yang, Y.; Jiao, L.; Liu, F. ViT-YOLO: Transformer-based YOLO for object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021.
40. Han, K.; Wang, Y.; Chen, H.; Chen, X.; Guo, J.; Liu, Z.; Tang, Y.; Xiao, A.; Xu, C.; Xu, Y.; et al. A Survey on Vision Transformer. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 87–110. [[CrossRef](#)] [[PubMed](#)]
41. Niu, Z.; Zhong, G.; Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **2021**, *452*, 48–62. [[CrossRef](#)]
42. Guo, M.-H.; Xu, T.-X.; Liu, J.-J.; Liu, Z.-N.; Jiang, P.-T.; Mu, T.-J.; Zhang, S.-H.; Martin, R.R.; Cheng, M.-M.; Hu, S.-M. Attention mechanisms in computer vision: A survey. *Comput. Vis. Media* **2022**, *8*, 331–368. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.