Generative Modelling of Temperature Field Evolution from Sparse Pressure Measurement

Shen Fang[1]; **Zeyu Li**[1]; **Qingfei Fu**[1,2]; **Wang Han**[1,2] **and Lijun Yang**[1,2]

[1] School of Astronautics, Beihang University, Beijing 100191, China

[2] National Key Laboratory of Aerospace Liquid Propulsion, Xi'an 710100, China

Keywords: Temperature Field Evolution, Generative modelling, Diffusion Transformer

Abstract. Measurement and modeling of temperature field evolution present critical challenges in aviation and aerospace engines due to harsh operating conditions, measurement errors, and the limited time resolution of traditional methods. In this study, we propose a novel generative modeling approach to predict temperature field evolution using high-fidelity, high-time-resolution pressure measurements. Specifically, we employ a Diffusion-Transformer (DiT) to learn a conditional generative model of temperature field dynamics based on sparse pressure signal series. This is achieved through latent diffusion modeling within a pre-trained autoencoder (AE) framework. To enhance modeling capability, frequency decomposition is applied to the temperature field. The proposed method is validated using a 2D shock wave reacting flow simulation dataset, demonstrating accurate predictions across the shock wave, post-shock, and pre-shock regions. The results highlight the efficacy of the DiT model for conditional generative modeling of physical field evolution, offering a promising solution for complex fluid dynamics applications.

1. Introduction

Combustion processes play a pivotal role in numerous industrial applications, from power generation and propulsion systems to chemical manufacturing, where the precise control of reaction dynamics is crucial for achieving optimal efficiency, ensuring operational safety, and reducing harmful emissions. Central to this control is the accurate measurement and inference of temperature fields, as they govern critical aspects such as reaction kinetics, heat transfer mechanisms, and the formation of pollutants. However, directly measuring temperature in combustion environments is exceptionally challenging due to the harsh conditions typically encountered, including extreme temperatures, rapid fluctuations, and the presence of chemically aggressive species. Conventional measurement techniques, such as thermocouples and pyrometers, are often constrained by limitations in spatial resolution, sluggish response times, and vulnerability to physical degradation under such conditions [1]. Moreover, intrusive measurement methods can disrupt the flow field, leading to measurement artifacts that skew the accuracy of the data [2]. These challenges highlight the urgent need for advanced, non-intrusive methodologies to infer temperature fields in combustion systems.

Pressure measurements present a robust and minimally intrusive alternative for monitoring combustion systems, offering distinct advantages over direct temperature measurement techniques. Modern pressure sensors, such as piezoelectric and piezoresistive transducers, are highly durable and capable of capturing high-frequency data with exceptional temporal resolution, all while minimizing disruption to the flow field [3]. This capability is particularly valuable in harsh combustion environments, where traditional temperature measurement methods often fail due to thermal degradation or spatial limitations. Furthermore, pressure fields are intrinsically linked to temperature fields through fundamental principles of fluid dynamics and thermodynamics, as described by the ideal gas law and the Navier-Stokes equations. These relationships suggest the potential to infer temperature fields from pressure measurements, provided that the underlying physical models are sufficiently accurate and validated [4]. Such a methodology could address the challenges associated with direct temperature measurement and open new avenues for real-time monitoring and control of combustion processes, thereby enhancing efficiency, safety, and emission reduction.

Recent advances in machine learning (ML) have significantly enhanced conventional optical and laser-based diagnostics in combustion research. ML techniques, particularly artificial neural networks, have

demonstrated remarkable capabilities in extracting complex and hidden correlations among different variables, enabling the prediction or reconstruction of one variable from others as inputs. For example, in turbulent combustion diagnostics, ML has been applied to reconstruct velocity fields from time-resolved hydroxyl (OH) radical images [5], generate planar distributions of soot particles from luminosity images [6], and enhance high-speed imaging of OH radicals [7]. These studies leverage the overlapping content embedded in simultaneous datasets, such as OH-PLIF and PIV fields, to reconstruct flame parameters [8]. Similarly, ML has been employed for flame pattern recognition, including instability detection in combustion systems [9] and classification of turbulent flame images [10]. Furthermore, deep learning models have been used to reconstruct planar distributions of species concentrations, such as CH radicals, from flame chemiluminescence in premixed turbulent flames [11]. These applications highlight the potential of ML to infer scalar fields, such as temperature or species concentrations, from related parameters. However, the use of ML to infer temperature fields from pressure measurements remains unexplored, presenting a significant opportunity for advancing combustion diagnostics.

To further expand the capabilities of machine learning in combustion diagnostics, this study introduces generative modeling for the prediction of temperature fields from sparse pressure measurement sequences. Unlike traditional methods that rely on direct correlations between inputs and outputs, generative models excel at capturing the underlying probabilistic distributions of complex systems, enabling the reconstruction of high-dimensional data from limited inputs [12-14]. Specifically, we employ a diffusion transformer framework, which leverages the strengths of diffusion models for high-fidelity data generation and transformers for capturing long-range dependencies in sequential data. This approach is applied to a shock wave flow dataset, where it demonstrates the ability to accurately reconstruct temperature fields from sparse pressure measurements. The results show that the proposed method achieves quantitative accuracy in predicting the temporal evolution of temperature fields, even in highly turbulent and non-linear combustion environments. By integrating these capabilities, the proposed method bridges a critical gap in combustion diagnostics, enabling the inference of temperature fields from pressure measurements—a task that remains unexplored in existing literature. This advancement not only enhances the diagnostic toolkit for combustion systems but also demonstrates the potential of generative models in addressing complex, multi-physics problems.

2. Methodology

2.1 Score-based generative model

Score-based generative models, widely known as diffusion models, have emerged as a powerful class of generative models in the machine learning community [12,15]. These models are currently considered the state of the art in tasks such as neural image synthesis[16] and language modelling[17]. The core idea of diffusion models is to construct a diffusion process $p_t(\mathbf{x})(t \in [0,T])$ that gradually transforms the data distribution p_{data} into a prior distribution p_{prior}, which is typically chosen as a standard Gaussian distribution. This transformation is achieved by learning the score function $\nabla_{\mathbf{x}} \log p_t(\mathbf{x})$ which represents the gradient of the log-likelihood of the data distribution at each step of the diffusion process. The generative process is realized through a denoising procedure, which can be formulated as the reverse of a stochastic differential equation (SDE) [18] given by

$$\mathrm{d}\mathbf{x} = [\mathbf{f}(\mathbf{x}, t) - g^2(t)\nabla_{\mathbf{x}} \log p_t(\mathbf{x})]\mathrm{d}t + g(t)\mathrm{d}\bar{\mathbf{w}} \tag{1}$$

where $\mathbf{f}(\mathbf{x}, t)$ is the drift function, $g(t)$ is the diffusion function and $\bar{\mathbf{w}}$ is a standard Wiener process with time flowing backward. Alternatively, a deterministic formulation of the generative process can be derived as a probability flow ordinary differential equation (ODE) [18], where the randomness is solely determined by the prior distribution. This ODE is expressed as

$$\mathrm{d}\mathbf{x} = -\dot{\sigma}(t)\sigma(t)\nabla_{\mathbf{x}} \log p_t\big(\mathbf{x}; \sigma(t)\big)\mathrm{d}t \tag{2}$$

where $\sigma(t)$ is the noise schedule function. By introducing an additional scale schedule $s(t)$, the ODE can be further generalised [19] to

$$\mathrm{d}\mathbf{x} = \left[\frac{\dot{s}(t)}{s(t)}\mathbf{x} - s^2(t)\dot{\sigma}(t)\sigma(t)\nabla_{\mathbf{x}}\log p_t\left(\frac{\mathbf{x}}{s(t)};\sigma(t)\right)\right]\mathrm{d}t \tag{3}$$

Recent work by [19] has significantly advanced diffusion models by elucidating key design choices, such as optimizing noise and scaling schedules and improving sampling methods. They introduced a second-order Heun method for solving the probability flow ODE, enhancing sampling accuracy and efficiency. These refinements have provided a clearer framework for designing and optimizing diffusion-based generative models.

2.2 Latent diffusion

Latent Diffusion Models (LDMs) [20] offer a computationally efficient framework for training diffusion models by operating in a compressed latent space rather than directly in high-dimensional pixel space. This two-stage approach begins with the training of an autoencoder, which learns to map high-resolution images $\mathbf{x}$ to a lower-dimensional latent representation $\mathbf{z} = \boldsymbol{E}(\mathbf{x})$, where $\boldsymbol{E}$ denotes the encoder. The corresponding decoder $\boldsymbol{D}$ reconstructs the image from the latent space, $\mathbf{x}' = \boldsymbol{D}(\mathbf{z})$. By compressing the input data into a latent space, LDMs significantly reduce the computational burden associated with high-resolution image generation.

The key innovation of LDMs lies in their application of the diffusion process to the latent space rather than the pixel space. This approach retains the theoretical foundations of diffusion models while drastically reducing computational costs, as the latent space is typically orders of magnitude smaller than the original pixel space. Additionally, LDMs benefit from the flexibility of the autoencoder framework, which allows for the use of pre-trained encoders and decoders, further enhancing their scalability and efficiency.

2.3 Diffusion transformer

Diffusion Transformers (DiTs) [14] introduce a transformer-based architecture as an alternative to the traditional U-Net backbone in diffusion models, offering significant advantages in scalability and global dependency modeling. Unlike U-Nets, which rely on convolutional operations to process spatial features hierarchically, DiTs operate on sequences of latent patches. This is achieved by patchifying the latent representation into non-overlapping patches, which are then treated as tokens and processed via self-attention mechanisms. This patch-based approach allows DiTs to capture long-range interactions across the entire latent space, a capability that is inherently limited in U-Nets due to their localized convolutional filters.

The patchification process introduces a trade-off between computational efficiency and model expressiveness. Smaller patch sizes increase the number of tokens, enabling finer-grained processing and improved performance at the cost of higher computational complexity. Conversely, larger patch sizes reduce computational overhead but may constrain the model's ability to capture fine details. By leveraging the transformer's ability to scale with increased depth, width, and token count, DiTs achieve state-of-the-art performance in terms of sample quality and computational efficiency, outperforming U-Net-based diffusion models.

2.4 Experiment settings

This study employs shock wave turbulent flow fields generated through Direct Numerical Simulation (DNS) combustion at two distinct Mach numbers: 0.1 and 0.3. For each Mach number condition, three randomly initialized flow fields are simulated, yielding a total of six sub-datasets. Among these, five sub-datasets are utilized for training, while the remaining one is designated for testing. Each sub-dataset comprises 576 frames, with each frame representing a scalar field of dimensions 512×512. To encapsulate the temporal dynamics of the flow field, sequential sets of five frames are treated as individual samples, capturing the evolution of the field over time. The primary focus of the dataset is the temperature field, with sparse pressure measurements serving as the conditioning input. Pressure conditions are sampled from consistent positions within the flow field across the entire dataset. Specifically, 50 measurement points are used for each frame, mimicking the real-world measurement of pressure in practical operating conditions. These pressure measurements are integrated into the DiT

architecture by embedding the conditioning information directly into the model via a cross-attention mechanism.

To augment the model's capacity to generate intricate details within the flow field, high-frequency components of the temperature field are extracted using a Gaussian Blur technique. These components are subsequently concatenated as additional channels alongside the original temperature field. Consequently, the input data assumes a shape of 5×3×512×512, where the three channels correspond to the original temperature field and its high-frequency components. The latent space for the latent diffusion process is derived from an autoencoder, for which we adopt a pretrained Deep Compression Autoencoder (DC-AE) from [21]. The autoencoder achieves a 64-time compression of the images, reducing the scalar field dimensions from 512×512 to 8×8 while increasing the number of channels from 3 to 128. This results in a latent representation with a shape of (128, 8, 8). By treating every five frames as a single sample, the input for the DiT model assumes a shape of (128, 5, 8, 8). The detailed structure and modules of the DiT model are illustrated in Figure 1. Further details regarding the DiT model parameters and training configurations are provided in Table 1.

Table 1. Default hyperparameters

	Parameter	Value
Network	Input size	(5,8,8)
	Input channel	128
	Patch size	(1,1,1)
	Hidden size	512
	Layers	12
	Heads	8
Training	Batch size	256
	Learning rate	1e-4
	Iterations	2e6
	Ema rate	0.999

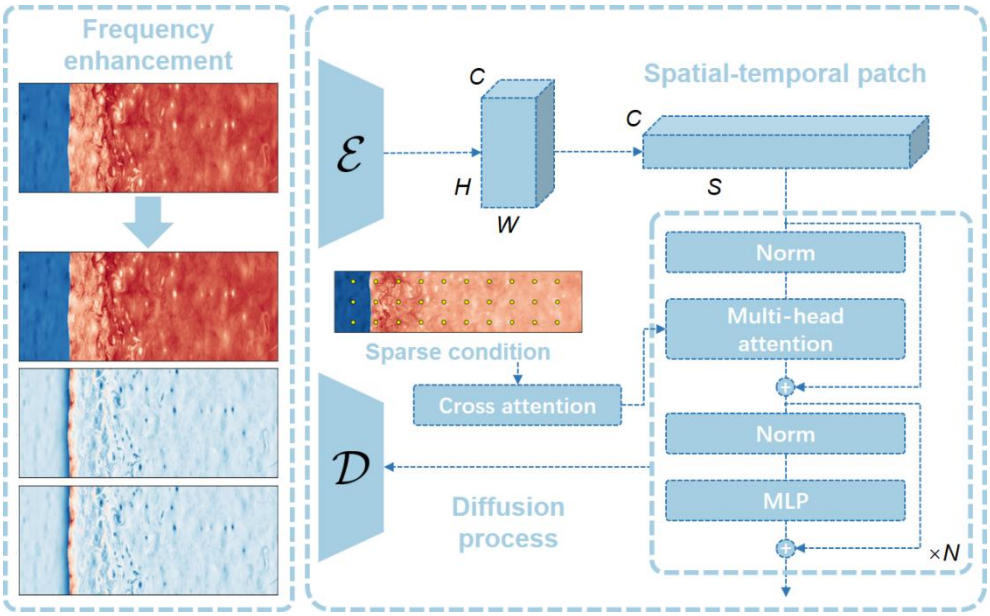

Figure 1. Model architecture

3. Results

3.1 Validation of autoencoder

The autoencoder employed in this study is first examined to validate its suitability for generating the latent space used in the diffusion process. As illustrated in Figure 2, the autoencoder achieves faithful

reconstruction of the temperature field, with the root mean square error (RMSE) being negligible. This performance is particularly noteworthy given that the pretrained autoencoder was originally trained on daily-life datasets such as ImageNet[22] and Mapillary Vistas[23], which do not include any data related to turbulent shock wave flow fields. Notably, the model was neither pretrained nor fine-tuned on such scientific datasets. The remarkable generalization capability of the autoencoder in encoding the latent space for this specialized task highlights the potential of leveraging pretrained models from traditional AI domains, such as image and language processing, for scientific machine learning applications. This finding underscores the versatility of pretrained models and suggests a promising avenue for cross-disciplinary research in scientific computing.

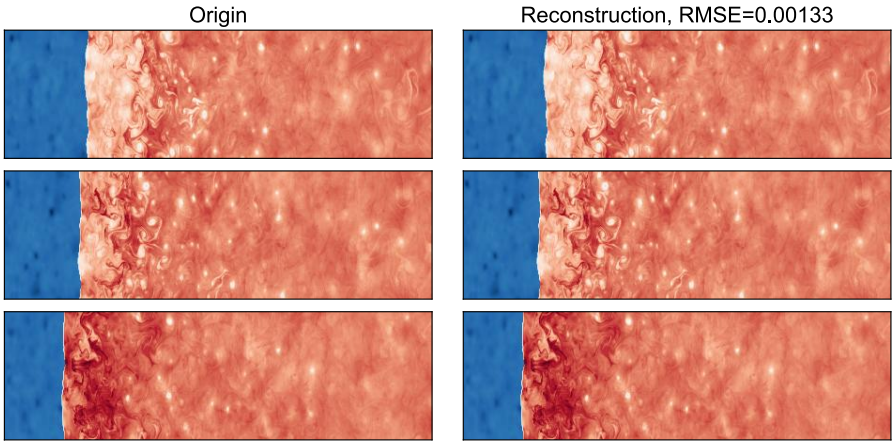

Figure 2. Autoencoder validation

3.2 Temperature field evolution prediction

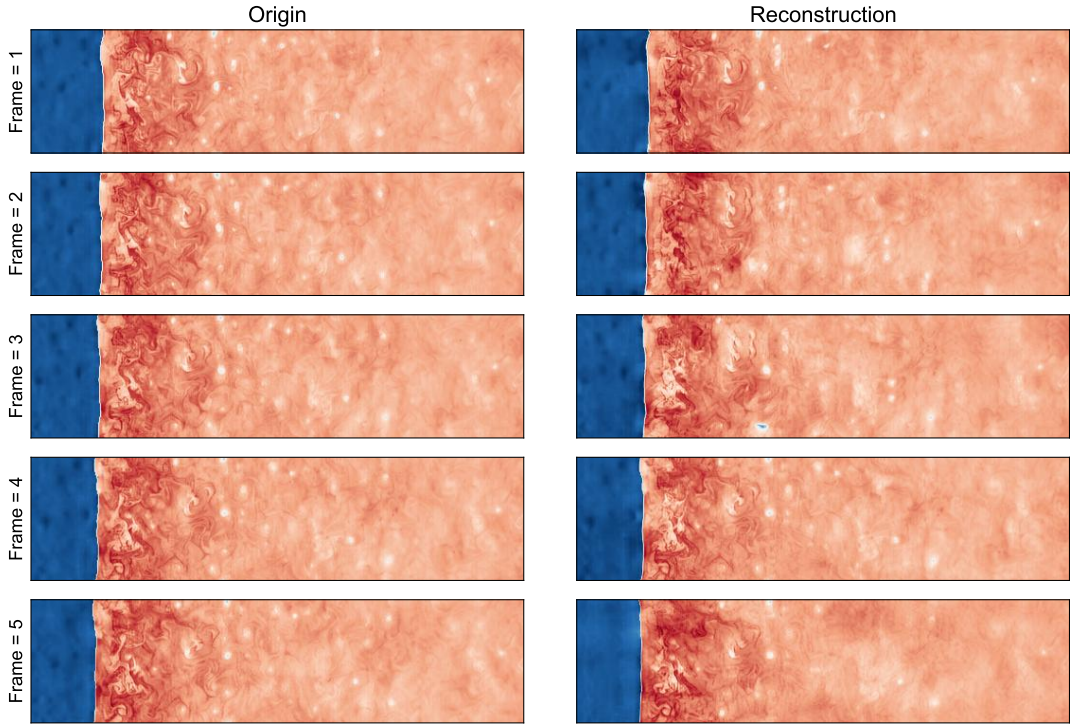

Figure 3. Representative origin and reconstructed temperature field evolution

Figure 3 illustrates the representative original and reconstructed temperature field evolution from the Diffusion Transformer (DiT) model, conditioned on sparse pressure measurements. The model effectively recovers the global temperature fields, accurately capturing the low-temperature region before the shock wave and the high-temperature region after the shock wave. Additionally, the higher

temperature near the shock wave, compared to the region closer to the outlet, is well described. In more detail, the model adeptly depicts the turbulent flow region just behind the shock wave, where the vortices are remarkably well-matched. Furthermore, the high-temperature small vortices, represented as white points in Figure 3, are also reconstructed by the model, albeit with some positional discrepancies. The successful reconstruction of such fine details, including the high-temperature singular points, demonstrates the DiT model's capability to learn the prior of the flow fields.

4. Conclusion

This study introduces a diffusion transformer framework for predicting temperature fields from sparse pressure measurements in turbulent combustion environments. The model accurately reconstructs temperature field evolution, even under highly non-linear conditions, demonstrating its ability to capture complex dynamics. By bridging a critical gap in combustion diagnostics, this work highlights the potential of generative models to solve multi-physics problems and advance scientific machine learning.

5. References

[1] Laser Diagnostics for Combustion Temperature and Species, CRC Press, 1996; eBook ISBN: 9781003077251. Author: Alan C. Eckbreth
[2] K. Kohse-Höinghaus, R. Barlow, M. Aldén, J. Wolfrum 'Combustion at the focus: laser diagnostics and control', P Combust Inst, https://doi.org/10.1016/j.proci.2004.08.274
[3] Springer Handbook of Experimental Fluid Mechanics, Springer, 2007; eBook ISBN: 9783540302995. Editors: Cameron Tropea, Alexander L. Yarin, John F. Foss
[4] Theoretical and Numerical Combustion (2nd ed.), R.T. Edwards Inc, 2005; eBook ISBN: 9781930217102. Authors: Thierry Poinsot, Denis Veynante
[5] S. Barwey, M. Hassanaly, V. Raman, A. Steinberg 'Using machine learning to construct velocity fields from OH-PLIF images', Combust Sci Technol, https://doi.org/10.1080/00102202.2019.1678379
[6] W. Zhang, X. Dong, C. Liu, et al. 'Generating planar distributions of soot particles from luminosity images in turbulent flames using deep learning', Appl Phys B, https://doi.org/10.1007/s00340-020-07571-9
[7] H. Guo, W. Zhang, X. Nie, et al. 'High-speed planar imaging of OH radicals in turbulent flames assisted by deep learning', Appl Phys B, https://doi.org/10.1007/s00340-021-07742-2
[8] S. Barwey, V. Raman, A. Steinberg 'Extracting information overlap in simultaneous OH-PLIF and PIV fields with neural networks', P Combust Inst, https://doi.org/10.1016/j.proci.2020.06.180
[9] T. Gangopadhyay, V. Ramanan, A. Akintayo, et al. '3D convolutional selective autoencoder for instability detection in combustion systems', Energy AI, https://doi.org/10.1016/j.egyai.2021.100067
[10] R. Roncancio, A. El Gamal, J. Gore 'Turbulent flame image classification using convolutional neural networks', Energy AI, https://doi.org/10.1016/j.egyai.2022.100193
[11] L. Han, Q. Gao, D. Zhang, et al. 'Deep neural network-based generation of planar CH distribution through flame chemiluminescence in premixed turbulent flame', Energy AI, https://doi.org/10.1016/j.egyai.2022.100221
[12] J. Ho, A. Jain, P. Abbeel 'Denoising Diffusion Probabilistic Models', Advances in Neural Information Processing Systems 34 (NeurIPS 2020), Virtual-only Conference, 6-12 December 2020
[13] T. Karras, M. Aittala, T. Aila, S. Laine 'Elucidating the Design Space of Diffusion-Based Generative Models', Advances in Neural Information Processing Systems 36 (NeurIPS 2023), New Orleans, USA, 10-16 December 2023
[14] W. Peebles, S. Xie 'Scalable Diffusion Models with Transformers', 2023 IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 1-6 October 2023
[15] Yang Song and Stefano Ermon 'Generative modeling by estimating gradients of the data distribution', Advances in Neural Information Processing Systems 33 (NeurIPS 2019), Vancouver, Canada, 8-14 December 2020
[16] P. Dhariwal and A. Q. Nichol 'Diffusion models beat GANs on image synthesis', Advances in Neural Information Processing Systems 35 (NeurIPS 2021), Virtual-only Conference, 6-14 December 2021
[17] Xiang Lisa Li, John Thickstun, Ishaan Gulrajani, et al. 'Diffusion-LM Improves Controllable Text Generation', Advances in Neural Information Processing Systems 36 (NeurIPS 2022), New Orleans, USA, 28 November-8 December 2022
[18] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, et al. 'Score-Based Generative Modeling through Stochastic Differential Equations', 9th International Conference on Learning Representations, Virtual-only Conference, 3-7 May 2021
[19] Tero Karras, Miika Aittala, Timo Aila, et al. 'Elucidating the Design Space of Diffusion-Based Generative Models', Advances in Neural Information Processing Systems 36 (NeurIPS 2022), New Orleans, USA, 28 November-8 December 2022
[20] R. Rombach, A. Blattmann, D. Lorenz, et al. 'High-resolution image synthesis with latent diffusion models', 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, USA, 18-24 June 2022
[21] Junyu Chen, Han Cai, Junsong Chen, et al. 'Deep Compression Autoencoder for Efficient High-Resolution Diffusion Models', 13th International Conference on Learning Representations, Singapore, Singapore, 24-28 April 2025
[22] J. Deng, W. Dong, R. Socher, et al. Imagenet: A large-scale hierarchical image database. 2009 IEEE conference on computer vision and pattern recognition (CVPR), Miami, USA, 22-24 June 2009
[23] G. Neuhold, T. Ollmann, S. Rota Bulo et al. The mapillary vistas dataset for semantic understanding of street scenes. 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22-29 October, 2017