

Semantic Loss as a System Metric for Imagery Transmission: DJSCC-Based Modeling and Resource Allocation

Ti Ti Nguyen, Vu Nguyen Ha, Thanh-Dung Le, Duc-Dung Tran, Hung Nguyen-Kha, Dinh-Hieu Tran, Carlos L. Marcos-Rojas, Raudel Cuiman, Juan C. Merlano-Duncan, and Symeon Chatzinotas

Abstract—Modern remote monitoring systems (RMS) increasingly rely on high-resolution imagery to support critical applications such as environmental monitoring, disaster response, and land-use analysis. Although these applications benefit from increasingly fine spatial resolution and higher temporal sampling rates of modern visual sensors, the resulting growth in imagery data volume imposes significant challenges on RMS constrained by limited bandwidth, power, and dynamic link conditions. To address these limitations, this paper investigates Deep Joint Source–Channel Coding (DJSCC) as an effective SemCom paradigm for the imagery transmission. We focus on two complementary aspects of semantic loss in DJSCC-based systems. First, a reconstruction-centric framework is evaluated by analyzing the semantic degradation of reconstructed images under varying compression ratios and channel signal-to-noise ratios (SNR), where semantic quality is quantified using reconstruction fidelity and perceptual similarity metrics. Second, a task-oriented framework is developed by integrating DJSCC with lightweight, application-specific models, with semantic quality measured directly through downstream task performance metrics rather than pixel-level fidelity. Based on extensive empirical analysis, we propose a unified semantic loss framework that captures both reconstruction-centric and task-oriented performance within a single model. This framework characterizes the implicit relationship between DJSCC compression, channel SNR, and semantic quality, offering actionable insights for the design of robust and efficient imagery transmission under resource-constrained satellite links. We then investigate a resource allocation framework for satellite systems that leverages the semantic loss model to maximize the minimum weighted worst-case semantic performance among ground users executing Earth observation imagery-based applications. Extensive numerical results validate the effectiveness of the proposed loss model and the associated resource management strategy.

Index Terms—Semantic communication, DJSCC, loss modeling, machine learning, remote monitoring.

I. INTRODUCTION

MODERN remote monitoring systems (RMS) increasingly rely on high-resolution imagery to support a

This work was funded by the Luxembourg National Research Fund (FNR), with granted SENTRY project corresponding to grant reference C23/IS/18073708/SENTRY.

T. Nguyen, V. N. Ha, D.-D. Tran, H. Nguyen-Kha, D.-H. Tran, C. L. Marcos-Rojas, R. Cuiman, J. C. Merlano-Duncan, S. Chatzinotas, are with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, 1855 Luxembourg Ville, Luxembourg. (e-mail: {titi.nguyen; vu-nguyen.ha; duc.tran; khahung.nguyen; hieu.tran-dinh; carlos.marcosrojas; raudel.cuiman; juan.duncan; Symeon.Chatzinotas}@uni.lu). T. D. Le is with Conrad Blucher Institute, Texas A&M University-Corpus Christi (e-mail: thanh-dung.le@tamucc.edu).

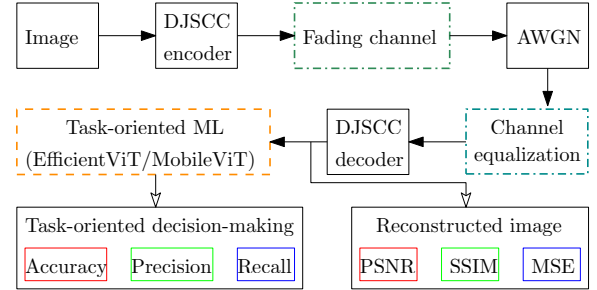


Fig. 1: Block diagram of SemCom with DJSCC.

wide range of applications, including smart city, smart home, environmental monitoring, disaster response, agriculture, and land-use analysis [1]. Recent advances in machine learning (ML) have enhanced RMS by enabling efficient analysis of complex datasets, improving accuracy and predictive insights [2], [3]. However, the growing resolution and acquisition frequency of remote sensors have led to an unprecedented increase in data volume, posing significant challenges for communication systems. For example, in earth observation (EO) missions, the transmission of high-resolution imagery is constrained by limited bandwidth, strict power budgets, and dynamically varying channel conditions. As a result, raw EO data often cannot be transmitted reliably or in a timely manner to ground stations, limiting the effectiveness of downstream EO applications [4]. In monitoring IoT schemes, a massive number of distributed sensing devices, including cameras, multispectral sensors, and embedded vision nodes, continuously capture high-resolution visual data for diverse applications under strict energy and bandwidth constraints. To address these challenges, semantic communication (SemCom) has emerged as a transformative paradigm, shifting traditional bit-oriented transmission toward conveying data meaning, thereby reducing overhead and latency in constrained environments [5]–[8].

Among SemCom approaches, deep joint source–channel coding (DJSCC) has shown strong potential by enabling end-to-end learned mappings that jointly optimize compression and transmission robustness over noisy channels [9]–[12]. However, deploying DJSCC in wireless systems introduces new challenges. Unlike conventional communication systems, whose performance can be optimized using well-defined metrics such as bit error rate or signal-to-noise ratio, SemCom performance depends on how transmission impairments affect application-level outcomes. Recent works [13]–[15] investigate semantic loss models; however, the implicit and highly

coupled relationships among DJSCC compression parameters, channel conditions, and semantic quality remain unexplored. Specifically, [13] models distortion in SemCom systems as the sum of two independent components, namely source distortion and channel distortion. The work in [14] adopts a generalized logistic function to approximate semantic similarity; however, this formulation cannot fully capture the two-dimensional characteristics of SemCom systems. In contrast, [15] adopts a relatively simple sum of exponential term to approximate the average data reconstruction error. Overall, the lack of tractable and expressive models complicates system design, particularly when balancing image reconstruction quality against downstream task performance.

To fill this gap, this work aims to investigate two complementary aspects of employing DJSCC for imagery transmission over wireless channels. First, a reconstruction-centric DJSCC framework is analyzed, where performance is measured in terms of the visual fidelity of reconstructed images under varying compression ratios (CR) and channel signal-to-noise ratios (SNRs). Second, a task-oriented DJSCC framework is developed by integrating DJSCC with lightweight, task-specific ML models, such as EfficientViT, MobileViT, deployed at computationally capable ground stations. In this setup, performance is evaluated using application-level metrics that reflect the utility of the received data for imagery transmission-related tasks, rather than pixel-level similarity. Reproducible codes are available in our shared repository¹.

Built on our prior work using the DVB-S2 standard [16], this work extends the analysis from traditional image compression to employing DJSCC techniques. An SemCom block diagram with DJSCC is illustrated in Fig. 1. In particular, leveraging extensive empirical evaluations on real-world data, this work proposes a unified data-driven semantic loss model that captures both reconstruction-centric and task-oriented performance within a common framework. Specifically, the EuroSAT dataset [17] is exploited for land use and land cover classification, based on which DJSCC performance with and without EfficientViT [18], MobileViT [19], with and without channel fading and channel equalization will be evaluated under varying source compression and channel conditions [20]. Explicitly, nine scenarios will be evaluated, which are:

- **DJSCC-AWGN**: Reconstructed image under an AWGN channel, without fading or equalization;
- **DJSCC-Fading-NoEqualization**: Reconstructed image under channel fading with AWGN, without equalization;
- **DJSCC-Fading-Equalization**: Reconstructed image under channel fading with AWGN, with equalization;
- **EfficientViT-DJSCC-AWGN**: Task-oriented decision-making using EfficientViT under an AWGN channel, without fading or equalization;
- **EfficientViT-DJSCC-NoEqualization**: Task-oriented decision-making using EfficientViT under channel fading with AWGN, without equalization;
- **EfficientViT-DJSCC-Equalization**: Task-oriented decision-making using EfficientViT under channel fading with AWGN, with equalization;

- **MobileViT-DJSCC-AWGN**: Task-oriented decision-making using MobileViT under an AWGN channel, without fading or equalization;
- **MobileViT-DJSCC-NoEqualization**: Task-oriented decision-making using MobileViT under channel fading with AWGN, without equalization;
- **MobileViT-DJSCC-Equalization**: Task-oriented decision-making using MobileViT under channel fading with AWGN, with equalization.

The proposed model empirically characterizes the relationship among DJSCC compression settings, channel SNR, and semantic effectiveness. This unified semantic performance framework provides practical guidance for designing robust, bandwidth-efficient, and task-aware imagery communication systems. These mathematical models are then employed to develop a resource allocation framework that jointly optimizes transmit power and semantic CRs under transmit power constraints to achieve various fairness-oriented objectives.

It is worth noting that channel equalization requires knowledge of the channel state information (CSI) at the receiver. Accurate CSI enables compensation for propagation-induced impairments, thereby mitigating signal distortions caused by channel fading. However, CSI may not always be available in all wireless systems, especially in satellite systems or systems with strict latency requirements. In such cases, channel equalization cannot be deployed. Even under these conditions, the DJSCC decoder can implicitly learn the channel distortions, and the end-to-end system performance is only slightly reduced compared to the case with equalization, as shown in Tables I-IX. The details of these tables are discussed in the following section.

Beyond robustness to imperfect or unavailable CSI, the semantic loss modeling framework developed in this work enables a systematic optimization of communication resources. While resource allocation has attracted considerable recent attention [21]–[25], existing studies lack an explicit and tractable semantic loss model, which limits their ability to achieve efficiently optimized solutions. By explicitly quantifying the relationship between DJSCC CR, channel SNR, and semantic performance, the proposed model provides a compact and optimization-ready representation of end-to-end semantic quality. This abstraction allows resource allocation problems, such as power and bandwidth assignment across multiple users with heterogeneous imagery-driven objectives, to be formulated directly in the semantic domain, rather than relying on conventional physical-layer metrics alone. In particular, the semantic loss model serves as a bridge between learning-based DJSCC transmission and system-level optimization, enabling the design of resource management strategies that explicitly target semantic performance fairness and robustness under dynamic wireless channel conditions. This perspective positions semantic loss as a first-class system metric, paving the way for cross-layer, semantic-aware wireless network design.

The remainder of this paper is organized as follows. Section II describes the imagery data acquisition process and the experimental setup. Section III presents the proposed curve-fitting-based semantic loss modeling methodology, detailing how reconstruction-centric and task-oriented semantic degr-

¹<https://github.com/ltdung/SnT-SENTRY>

dation is jointly characterized. In Section IV, the developed semantic loss model is leveraged to formulate and solve a semantic-aware resource allocation problem for imagery missions, illustrating its role in optimizing power and bandwidth allocation under heterogeneous user and application requirements. Finally, Section V concludes the paper.

II. DATA ACQUISITION FOR SEMANTIC-LOSS MODELING

A. DJSCC-based Imagery Transmission

In this paper, we explore the performance of employing DJSCC for imagery transmission in an RMS [9]. Specifically, images captured by distributed imaging platforms (i.e., the original dataset) are transmitted to remote processing servers for downstream analysis.

1) *Deep Learning-Based Joint Source-Channel Coding*: A DJSCC scheme models the end-to-end communication system as a deep neural network (DNN) based autoencoder, where source compression and channel coding are learned jointly [9]. Specifically, an encoder $f_\theta : \mathbb{R}^n \rightarrow \mathbb{C}^k$, parameterized by a convolutional neural network (CNN), maps the source signal $\mathbf{x}$ directly to complex-valued channel symbols $\mathbf{z}$ without explicit quantization or channel coding. The encoder output is normalized to satisfy an average transmit power constraint

$$\mathbf{z} = \frac{\sqrt{kP} \tilde{\mathbf{z}}}{\sqrt{\tilde{\mathbf{z}}^* \tilde{\mathbf{z}}}}, \quad (1)$$

where $\tilde{\mathbf{z}}$ denotes the unnormalized CNN output and P is the power budget. The communication channel is incorporated into the DNN as a non-trainable differentiable layer

$$\hat{\mathbf{z}} = \eta(\mathbf{z}), \quad (2)$$

enabling end-to-end training via backpropagation. For example, $\eta(\cdot)$ models an AWGN channel as $\hat{\mathbf{z}} = \mathbf{z} + \mathbf{n}$, or a slow fading channel as $\hat{\mathbf{z}} = h\mathbf{z} + \mathbf{n}$. We assume that the fading coefficient h remains constant within each batch transmission.

At the receiver, a CNN-based decoder $g_\phi : \mathbb{C}^k \rightarrow \mathbb{R}^n$ reconstructs the source signal $\hat{\mathbf{x}}$ either directly from the received signal $\hat{\mathbf{z}}$ (**without equalization**) or from the **equalized signal** $h^{-1}\hat{\mathbf{z}}$, depending on whether channel equalization is employed. The encoder and decoder parameters (θ, ϕ) are jointly optimized by minimizing the expected distortion

$$(\theta^*, \phi^*) = \arg \min_{\theta, \phi} \mathbb{E}_{\mathbf{x}}[d(\mathbf{x}, \hat{\mathbf{x}})], \quad (3)$$

where $d(\cdot)$ denotes a distortion metric (e.g., MSE). Using DNNs, DJSCC learns compact source embeddings and robust channel mappings, resulting in graceful performance degradation as channel quality decreases, eliminating the cliff effect typical of separation-based digital transmission schemes.

2) *Compression Ratio and Channel Quality*: Thanks to the DJSCC-based transmission, the images are then reconstructed at the ground station. Image classification is then employed as a representative downstream application, using these reconstructed images for performance evaluation. Unlike traditional communication systems that rely on separate source coding (e.g., JPEG2000) and channel coding (e.g., LDPC or Turbo codes), DJSCC leverages deep neural networks to jointly learn both encoding and decoding functionalities in

TABLE I: DJSCC-AWGN EMPIRICAL RESULTS.

$\gamma \backslash \rho$	2	4	6	8	12
PSNR					
8 dB	35.1659	34.5505	34.0580	33.6110	33.2048
4 dB	34.6355	33.7960	33.1808	32.7066	32.1396
0 dB	33.9541	32.9471	32.2797	31.6707	30.9957
-1 dB	33.6718	32.6668	31.9636	31.4396	30.7259
-2 dB	33.3420	32.4279	31.7131	31.0607	30.3161
-3 dB	33.0512	32.0707	31.3065	30.7091	29.9191
-4 dB	32.7076	31.7420	30.9968	30.3004	29.4664
-5 dB	32.4621	31.3836	30.5894	29.9658	29.0585
-6 dB	32.0890	31.0723	30.2680	29.5925	28.6726
-7 dB	31.8650	30.7567	29.9635	29.2635	28.3096
SSIM					
8 dB	93.9029	93.0065	92.0605	90.8419	89.3904
4 dB	92.8318	91.0260	89.9122	88.5497	86.9133
0 dB	91.4196	88.2488	86.1295	84.2376	81.8966
-1 dB	90.7393	87.6952	85.3222	83.4208	80.6235
-2 dB	90.2920	86.7390	84.0706	81.7849	78.9081
-3 dB	89.2889	85.2669	82.4933	80.2096	76.8448
-4 dB	88.2929	83.8135	80.8698	78.0683	74.5515
-5 dB	86.9725	82.3352	78.8621	76.3074	72.1772
-6 dB	85.5012	80.6746	77.0950	74.1351	69.7947
-7 dB	84.3367	78.9294	75.2783	72.1513	67.5156
MSE					
8 dB	30.1747	35.2691	40.1839	47.5344	58.4974
4 dB	36.4795	48.8558	55.1101	63.4826	76.4652
0 dB	44.4293	64.2171	77.8224	88.3959	106.3604
-1 dB	49.6283	65.6896	78.7179	90.9191	111.2500
-2 dB	50.5202	71.9030	86.6523	102.2033	123.2206
-3 dB	57.5215	80.8606	97.6941	111.2566	134.7072
-4 dB	63.0393	89.4421	106.4504	126.8159	151.5233
-5 dB	70.4309	99.1093	122.2410	136.6730	167.8200
-6 dB	80.5933	110.0527	134.3242	153.9908	185.4238
-7 dB	87.7364	123.5263	147.7795	170.0484	204.4302

TABLE II: DJSCC-FADING-NOEQUALIZATION EMPIRICAL RESULTS.

$\gamma \backslash \rho$	2	4	6	8	12
PSNR					
8 dB	32.8278	31.6357	31.0843	30.6777	30.1185
4 dB	31.8538	31.0293	30.2306	29.5990	28.5822
0 dB	31.2313	30.1468	29.2175	28.6782	27.2661
-1 dB	31.0403	30.0557	29.0528	28.4770	27.1696
-2 dB	30.8645	29.4920	28.6727	28.2349	26.7591
-3 dB	30.3631	29.1563	28.4265	27.9356	26.4024
-4 dB	30.0805	28.8171	28.0543	27.6021	26.0992
-5 dB	29.6458	28.4563	27.8734	27.2788	25.8171
-6 dB	29.1549	28.1840	27.3419	26.9251	25.5101
-7 dB	28.9298	27.9086	27.2038	26.6510	25.2786
SSIM					
8 dB	88.5785	84.9228	84.3280	83.0973	81.0289
4 dB	85.3523	82.7210	81.5555	79.3781	76.2843
0 dB	82.8903	79.4061	76.3312	74.6789	71.8988
-1 dB	82.5441	78.5874	76.0233	73.6953	71.2417
-2 dB	81.7098	76.6176	74.1708	72.4531	69.9069
-3 dB	79.9276	75.4835	72.6296	71.1069	68.6380
-4 dB	78.7782	73.4504	71.3053	69.9589	67.2303
-5 dB	76.9180	72.4103	70.1662	68.4572	65.9218
-6 dB	74.7663	71.0963	68.6249	67.0551	64.4474
-7 dB	73.9827	70.0346	67.8215	65.7864	63.3263
MSE					
8 dB	67.2714	88.3280	97.5750	106.3747	121.4941
4 dB	86.5259	105.8419	120.8968	137.3567	168.1911
0 dB	104.4137	134.0866	160.0710	177.3695	223.5098
-1 dB	107.7057	141.2495	165.7993	188.3001	232.3153
-2 dB	116.4742	154.7739	179.9386	199.2700	251.9488
-3 dB	127.6451	167.4642	195.1889	213.9688	270.9011
-4 dB	137.9119	182.2894	208.9204	230.3854	291.5352
-5 dB	152.7359	196.7462	222.7411	247.8974	313.0898
-6 dB	168.2614	211.6015	242.8162	268.8060	337.4685
-7 dB	179.3620	225.6840	256.7700	283.7463	355.4150

a single end-to-end model. This enables DJSCC to more effectively preserve semantic information under bandwidth and channel constraints, particularly in non-ideal communication environments where the separation theorem does not hold.

As illustrated in Fig. 1, performance degradation in Sem-Com systems using DJSCC stems from two primary sources:

TABLE III: DJSCC-FADING-EQUALIZATION EMPIRICAL RESULTS.

$\gamma \backslash \rho$	2	4	6	8	12
PSNR					
8 dB	33.9384	33.3300	32.8683	32.5410	32.0906
4 dB	33.3318	32.5965	31.9952	31.5435	30.8335
0 dB	32.7284	31.8571	31.1387	30.6009	29.9857
-1 dB	32.4812	31.5892	30.8746	30.4295	29.7126
-2 dB	32.3080	31.4187	30.7111	30.2647	29.4402
-3 dB	31.9900	31.1690	30.4321	29.8926	29.1331
-4 dB	31.6868	30.7235	30.1232	29.5294	28.8635
-5 dB	31.3953	30.5197	29.6913	29.3199	28.6206
-6 dB	31.0941	30.2406	29.4422	29.0115	28.2761
-7 dB	30.8067	29.8980	29.1279	28.7082	27.9718
SSIM					
8 dB	90.8341	89.5699	88.2646	87.0535	85.6766
4 dB	89.2942	87.2912	85.5536	84.0222	81.8244
0 dB	87.1479	84.2678	82.1284	80.3020	78.4251
-1 dB	86.8683	83.7331	80.7949	79.2379	77.8826
-2 dB	86.2645	82.9135	80.3370	78.8700	76.3904
-3 dB	85.0750	82.0862	79.3407	77.4169	74.9171
-4 dB	84.0116	80.2567	78.0718	75.9133	73.6533
-5 dB	82.9476	79.4637	75.9334	74.9501	72.9063
-6 dB	81.7486	78.2999	74.7741	73.4959	71.2505
-7 dB	80.5794	76.5724	73.3504	72.1503	69.8297
MSE					
8 dB	51.6590	61.4559	72.0797	77.8509	85.5910
4 dB	61.9978	75.1716	88.0371	96.6932	112.3199
0 dB	76.0534	93.4839	109.7121	120.8245	137.9110
-1 dB	80.0550	99.1393	118.8697	129.1739	146.2209
-2 dB	84.4695	105.3199	124.2906	134.5358	156.7880
-3 dB	91.4935	111.8339	131.4145	144.8526	166.5482
-4 dB	98.6510	123.0529	140.7475	156.0209	176.0808
-5 dB	105.2676	128.9234	153.2908	163.4044	185.1441
-6 dB	113.6402	138.0037	162.5876	175.2017	198.5229
-7 dB	121.3990	149.5840	174.7141	186.2992	212.2693

TABLE IV: EMPIRICAL TASK-LEVEL RESULTS (%) USING EFFICIENTViT-DJSCC-AWGN.

$\gamma \backslash \rho$	2	4	6	8	12
Accuracy					
8 dB	97.7051	97.2967	96.9067	96.5121	95.8632
4 dB	97.2928	96.5092	96.0357	95.3670	94.3556
0 dB	97.0890	95.6177	94.5010	93.4146	91.3953
-1 dB	96.8802	95.3873	93.9642	92.7888	90.1026
-2 dB	96.6108	94.8307	93.1518	91.5226	88.5643
-3 dB	96.0471	93.7941	92.0042	90.1831	86.5272
-4 dB	95.4615	92.9421	90.5939	88.3597	84.3994
-5 dB	94.7738	91.9048	89.0595	86.6307	81.8966
-6 dB	93.9276	90.6833	87.5220	84.6269	79.1926
Precision					
8 dB	97.6660	97.3024	96.9487	96.5824	95.9765
4 dB	97.2554	96.5211	96.0831	95.4476	94.4837
0 dB	96.8145	95.6488	94.5692	93.5293	91.5745
-1 dB	96.8520	95.3238	94.0368	92.8982	90.2869
-2 dB	96.5923	94.8722	93.2397	91.6484	88.7577
-3 dB	96.0325	93.8358	92.1079	90.3163	86.7739
-4 dB	95.4552	92.9980	90.7076	88.5351	84.6581
-5 dB	94.7731	91.9842	89.1803	86.8053	82.1347
-6 dB	93.9442	90.7627	87.6891	84.8078	79.4080
Recall					
8 dB	97.6127	97.2508	96.8902	96.5088	95.8397
4 dB	97.2018	96.4655	96.0222	95.3674	94.3372
0 dB	96.7542	95.5818	94.4970	93.4260	91.3905
-1 dB	96.7982	95.1904	93.9644	92.8050	90.1037
-2 dB	96.5321	94.8024	93.1568	91.5443	88.5721
-3 dB	95.9718	93.7701	92.0142	90.2106	86.5420
-4 dB	95.3896	92.9224	90.6089	88.3929	84.4213
-5 dB	94.7052	91.8891	89.0794	86.6693	81.9252
-6 dB	93.8619	90.6713	87.5462	84.6706	79.2274

(1) source distortion introduced by the encoder–decoder compression process, and (2) channel distortion caused by noisy wireless transmission. To analyze and quantify these effects, we introduce two key parameters: (a) the CR ρ , defined as the ratio between the dimensionality of the original image tensor and that of the encoded feature tensor, and (b) the SNR γ , which characterizes the physical channel quality.

TABLE V: EMPIRICAL TASK-LEVEL RESULTS (%) USING EFFICIENTViT-DJSCC-NOEQUALIZATION.

$\gamma \backslash \rho$	2	4	6	8	12
Accuracy					
8 dB	94.5553	92.7561	90.9259	89.9773	87.6751
4 dB	92.7777	90.2874	87.6827	85.7330	81.0465
0 dB	90.7348	86.5646	82.5043	79.8255	71.3141
-1 dB	90.2404	85.6120	81.8096	78.1254	69.6978
-2 dB	88.9614	83.6813	79.5658	76.3626	66.3860
-3 dB	87.4329	81.9008	77.8306	73.8576	63.1308
-4 dB	86.1401	80.0181	75.4439	71.3704	60.0408
-5 dB	84.4072	78.0762	73.9175	68.9317	57.0140
-6 dB	82.2931	76.2995	70.2380	66.0399	54.0941
Precision					
8 dB	94.5790	93.0340	91.0243	90.1677	87.8120
4 dB	92.7336	90.5440	87.7422	85.9408	81.3028
0 dB	90.8507	86.8562	82.6536	80.1325	71.5279
-1 dB	90.3646	85.8935	81.9656	78.4739	69.9724
-2 dB	89.1570	84.0057	79.7475	76.7030	66.5640
-3 dB	87.6594	82.2573	78.0257	74.2451	63.3111
-4 dB	86.3795	80.3572	75.6470	71.7074	60.2344
-5 dB	84.6579	78.3944	74.0791	69.2380	57.1383
-6 dB	82.5926	76.5746	70.4210	66.3860	54.2441
-7 dB	82.5926	76.5746	70.4210	66.3860	54.2441
Recall					
8 dB	94.5702	92.7612	90.9041	89.9662	87.6710
4 dB	92.7940	90.3065	87.6805	85.7414	81.0433
0 dB	90.7135	86.5701	82.5047	79.8468	71.3235
-1 dB	90.2096	85.6049	81.7969	78.1338	69.6948
-2 dB	88.9335	83.6714	79.5484	76.3664	66.3797
-3 dB	87.4174	81.8954	77.8147	73.8624	63.1265
-4 dB	86.1416	80.0206	75.4320	71.3777	60.0383
-5 dB	84.4267	78.0868	73.9091	68.9403	57.0102
-6 dB	82.3291	76.3170	70.2315	66.0474	54.0851

TABLE VI: EMPIRICAL TASK-LEVEL RESULTS (%) USING EFFICIENTViT-DJSCC-EQUALIZATION.

$\gamma \backslash \rho$	2	4	6	8	12
Accuracy					
8 dB	95.5818	94.8208	94.1603	93.2651	91.9036
4 dB	94.6724	93.4134	91.9142	90.8123	88.4001
0 dB	92.8195	91.4465	88.7674	87.3265	84.5894
-1 dB	92.5499	90.3090	88.1692	86.6755	84.1675
-2 dB	92.4411	89.8008	87.8225	86.5249	82.2638
-3 dB	91.1786	89.3929	87.2953	84.6029	80.0349
-4 dB	90.4796	87.4484	85.4986	82.4671	78.4379
-5 dB	89.8608	87.3659	83.9502	82.3364	78.3606
-6 dB	89.0009	86.4918	83.3126	80.9549	75.2007
Precision					
8 dB	95.7303	94.9159	94.1751	93.3534	92.0139
4 dB	94.6232	93.3805	92.0482	90.9508	88.6613
0 dB	92.9957	91.3869	89.1810	87.6327	84.8951
-1 dB	92.6577	90.5254	88.5642	86.9352	84.1680
-2 dB	92.3404	89.9362	87.9571	86.4120	82.5996
-3 dB	91.3850	89.3161	87.1738	84.7914	80.6813
-4 dB	90.6540	87.8850	85.7404	83.0837	79.0896
-5 dB	89.9721	87.3658	84.2085	82.4121	78.1791
-6 dB	89.2081	86.5244	83.3739	81.0693	75.8026
Recall					
8 dB	95.6895	94.8554	94.0994	93.2550	91.8870
4 dB	94.5669	93.3021	91.9405	90.7991	88.4847
0 dB	92.9076	91.2571	89.0137	87.4449	84.5549
-1 dB	92.5901	90.4249	88.3811	86.7267	83.8653
-2 dB	92.2658	89.8126	87.7689	86.1389	82.2741
-3 dB	91.2883	89.1689	86.9765	84.5630	80.3804
-4 dB	90.5477	87.7117	85.5006	82.8252	78.7247
-5 dB	89.8363	87.1786	83.8942	82.0855	77.8180
-6 dB	89.0200	86.2729	83.0226	80.7551	75.4114

B. Original Dataset and DJSCC Implementation

The public benchmark for land use and land cover classification dataset, EuroSAT [17], will be used in this study. This is a large-scale benchmark dataset designed explicitly for land use and land cover classification derived from Sentinel-2 satellite imagery. It comprises 27,000 geo-referenced labeled images, each measuring 64x64 pixels and spanning 13 spectral bands. The dataset is categorized into 10 classes. Each class

TABLE VII: EMPIRICAL TASK-LEVEL RESULTS (%) USING MOBILEViT-DJSCC-AWGN.

$\gamma \backslash \rho$	2	4	6	8	12
Accuracy					
8 dB	97.8441	97.3829	97.1174	96.6812	96.0835
4 dB	97.4090	96.6031	96.0855	95.4389	94.5822
0 dB	97.2050	95.7667	94.7202	93.6903	91.5903
-1 dB	97.0881	95.4577	94.3503	93.0829	90.5369
-2 dB	96.8063	95.0773	93.5242	91.8676	88.8861
-3 dB	96.2819	94.0886	92.3881	90.5320	86.8269
-4 dB	95.6882	93.2360	90.9479	88.6295	84.7518
-5 dB	94.9958	92.2845	89.4833	86.9015	82.3349
-6 dB	94.1865	90.9417	87.8287	84.8275	79.5618
Precision					
8 dB	97.9522	97.6240	97.3684	96.9726	96.4163
4 dB	97.5213	96.7996	96.3398	95.7453	94.9330
0 dB	97.3051	95.9846	94.9993	94.0245	91.9720
-1 dB	97.1173	95.6822	94.6375	93.4353	90.9514
-2 dB	96.7966	95.3059	93.8300	92.2258	89.3304
-3 dB	96.4167	94.3194	92.7177	90.9164	87.3391
-4 dB	95.8318	93.4859	91.2816	89.0563	85.2850
-5 dB	95.1489	92.5476	89.8410	87.3395	82.8569
-6 dB	94.3493	91.2323	88.2083	85.2838	80.0850
Recall					
8 dB	97.9404	97.5430	97.3255	96.9221	96.3468
4 dB	97.5065	96.7654	96.2966	95.6837	94.8508
0 dB	97.2852	95.9368	94.9418	93.9478	91.8760
-1 dB	97.0942	95.6571	94.5769	93.3467	90.8309
-2 dB	96.7713	95.2558	93.7566	92.1386	89.1900
-3 dB	96.3935	94.2719	92.6268	90.8108	87.1415
-4 dB	95.8031	93.4241	91.1931	88.9164	85.0776
-5 dB	95.1137	92.4772	89.7348	87.1962	82.6714
-6 dB	94.3071	91.1388	88.0860	85.1294	79.9085

TABLE VIII: EMPIRICAL TASK-LEVEL RESULTS (%) USING MOBILEViT-DJSCC-NOEQUALIZATION.

$\gamma \backslash \rho$	2	4	6	8	12
Accuracy					
8 dB	94.7695	93.7131	92.2480	90.9378	88.3134
4 dB	93.2446	91.2728	88.6989	86.5886	82.0078
0 dB	91.2828	87.2817	83.5941	80.6313	73.0436
-1 dB	90.8352	86.5283	82.8098	79.5048	71.1266
-2 dB	89.7018	84.5281	80.9423	77.5804	68.1776
-3 dB	88.2274	82.9296	79.0996	75.1111	64.8960
-4 dB	87.0729	81.0804	76.9372	72.7191	61.6380
-5 dB	85.4819	79.2560	75.3775	70.2500	58.3843
-6 dB	83.5140	77.6142	71.9543	67.5389	55.3149
Precision					
8 dB	94.8720	93.8582	92.1675	91.1341	88.5527
4 dB	93.2612	91.3706	88.6246	86.8107	82.1931
0 dB	91.2556	87.4901	83.6549	80.9487	73.3133
-1 dB	90.8577	86.7332	82.9288	79.8774	71.4113
-2 dB	89.7654	84.8242	81.0323	78.0031	68.4468
-3 dB	88.3344	83.2498	79.2777	75.6069	65.1110
-4 dB	87.2211	81.4590	77.0952	73.1423	61.8647
-5 dB	85.6676	79.7117	75.5904	70.6562	58.5241
-6 dB	83.7678	78.0156	72.1905	68.0172	55.5107
Recall					
8 dB	94.8604	93.7256	92.0438	91.0071	88.4411
4 dB	93.2152	91.2652	88.4887	86.6296	82.0015
0 dB	91.1907	87.3069	83.4273	80.6901	72.9368
-1 dB	90.7645	86.5826	82.6748	79.5960	71.0500
-2 dB	89.6592	84.6176	80.8469	77.7151	68.1521
-3 dB	88.2082	83.0480	79.0375	75.2833	64.9181
-4 dB	87.0594	81.2067	76.8842	72.9015	61.6729
-5 dB	85.4502	79.3620	75.3008	70.4048	58.3816
-6 dB	83.4440	77.6761	71.8260	67.6333	55.2299

contains 2000–3000 images, including industrial buildings, residential buildings, annual crops, permanent crops, rivers, seas and lakes, Herbaceous vegetation, highways, pastures, and forests across Europe. Due to its compact image size and diverse class representation, EuroSAT is particularly well-suited for developing and evaluating DL models intended for onboard satellite processing. This makes it a valuable resource for applications such as real-time environmental monitoring, disaster response, and precision agriculture.

TABLE IX: EMPIRICAL TASK-LEVEL RESULTS (%) USING MOBILEViT-DJSCC-EQUALIZATION.

$\gamma \backslash \rho$	2	4	6	8	12
Accuracy					
8 dB	95.7871	95.2586	94.3309	93.4976	92.3707
4 dB	94.9150	93.5868	92.3554	91.1138	88.6942
0 dB	93.4492	91.7509	89.3657	87.6159	85.2276
-1 dB	93.0245	90.7468	88.9568	87.3724	84.8626
-2 dB	92.7745	90.3300	88.3764	86.9333	83.2504
-3 dB	91.6789	90.0184	87.8643	85.2713	81.3254
-4 dB	90.9837	88.1073	86.0382	83.5436	79.5804
-5 dB	90.3823	87.8429	84.3864	82.9707	78.9308
-6 dB	89.7204	87.1362	83.5215	81.5263	76.3317
Precision					
8 dB	96.0107	95.3219	94.4777	93.7267	92.5757
4 dB	94.9696	93.6910	92.5113	91.3911	88.9654
0 dB	93.5750	91.8668	89.7234	88.0550	85.5684
-1 dB	93.2197	91.0888	89.3359	87.7182	85.2008
-2 dB	92.9386	90.6284	88.7361	87.2160	83.8849
-3 dB	92.0468	90.1837	88.0833	85.7310	82.1531
-4 dB	91.3550	88.6211	86.5162	84.1867	80.5147
-5 dB	90.7102	88.1373	85.0698	83.2951	79.4666
-6 dB	90.0238	87.3144	84.0595	81.8531	77.0824
Recall					
8 dB	95.9953	95.2589	94.3901	93.6226	92.4527
4 dB	94.9304	93.6170	92.4163	91.2057	88.7468
0 dB	93.5339	91.7971	89.6292	87.9055	85.2903
-1 dB	93.1616	90.9889	89.1791	87.5444	84.8874
-2 dB	92.8570	90.5163	88.5496	86.9754	83.5353
-3 dB	91.9371	90.0546	87.8951	85.5040	81.8330
-4 dB	91.2403	88.4765	86.3046	83.9138	80.1233
-5 dB	90.5767	87.9639	84.7917	83.0442	79.0548
-6 dB	89.8615	87.1457	83.7909	81.5763	76.6466

To evaluate SemCom performance, we consider the Rayleigh fading and perfect CSI for channel equalization [20]. The batch size is 32, and the fading coefficient remains constant within each batch transmission. Furthermore, we implement the DJSCC framework proposed in [9] under different CRs ρ and SNR γ are considered. Specifically, we consider $\rho \in \{2, 4, 6, 8, 12\}$ and $\gamma \in \{-6, -5, -4, -3, -2, -1, 0, 4, 8\}$ (dB). Furthermore, motivated by recent findings in [26], which identify EfficientViT and MobileViT as a highly effective Vision Transformer for image classification due to its strong trade-off between performance and computational efficiency, we adopt EfficientViT and MobileViT as the downstream classifier in our evaluation. Accordingly, for each combination of CR ρ and SNR γ , we apply the DJSCC framework with and without EfficientViT or MobileViT. Our empirical results (shown in Tables I, II, III, IV, V, VI, VII, VIII, and IX) confirm that both ρ and γ jointly influence semantic performance metrics, including both reconstruction fidelity and task-oriented accuracy. Note that *peak signal-to-noise ratio (PSNR)*, *mean squared error (MSE)*, and *structural similarity index measure (SSIM)* are used to assess image reconstruction quality [13], whereas *accuracy*, *precision*, and *recall* are considered for evaluating task performance [26]. For convenience, the term semantic loss may refer to any of these metrics, MSE, PSNR, SSIM, accuracy, precision, or recall, depending on the performance objective under consideration.

III. METHODOLOGY - CURVE FITTING-BASED MODELING

To the best of our knowledge, there is no theoretical result in the literature for evaluating the semantic loss due to wireless image transmission. This motivates us to employ a practical data-fitting approach for modeling the semantic loss as a non-linear function of ρ and γ . Generally, choosing data-fitting

$$\xi(\rho, \gamma) = \mu_0 + \sum_{i=1}^{N_c} \left(\mu_{1,i} + \frac{\mu_{2,i}}{1 + \exp(-\mu_{3,i}\gamma - \mu_{4,i})} \right) (\exp(\mu_{5,i}\rho) + \mu_{6,i}\rho). \quad (4)$$

Algorithm 1 GRADIENT DESCENT FOR SEMCOM MODELING

```

1: Initialize: Set parameters  $\mu_i$  and learning rates  $\alpha_i, i = 0 : 6$ ,
   number of terms  $N_c$ , number of iterations  $N$ , measured data
   matrix  $\mathbf{Y}$  and set of control parameters  $\rho$  and  $\gamma$ .
2: Initialize gradients:  $\text{Grad}_{\mu_0} \leftarrow 0$ , and  $\text{Grad}_{\mu_i} \leftarrow \mathbf{0}$ , for  $i = 1 : 6$ .
3: for iter = 1 to  $N$  do
4:   Compute  $\xi$  as defined in (4).
5:   Compute error matrix:  $\Psi \leftarrow \mathbf{Y} - \xi$ .
6:   for  $i = 1$  to  $\text{length}(s)$  do
7:     for  $j = 1$  to  $\text{length}(q)$  do
8:       for  $k = 1$  to  $N_c$  do
9:         Compute  $\sigma_k \leftarrow \frac{1}{1 + \exp(-\mu_{3,k}\gamma_i - \mu_{4,k})}$ .
10:        Compute  $\eta_k \leftarrow \exp(\mu_{5,k}\rho_j)$ .
11:        Compute  $\beta_k \leftarrow \eta_k + \mu_{6,k}\rho_j$ .
12:        Update gradients: (where  $\epsilon_{ij} = [\Psi]_{i,j}$ )
13:         $\text{Grad}_{\mu_{1,k}} \leftarrow \text{Grad}_{\mu_{1,k}} - 2\epsilon_{ij}\beta_k$ .
14:         $\text{Grad}_{\mu_{2,k}} \leftarrow \text{Grad}_{\mu_{2,k}} - 2\epsilon_{ij}\beta_k\sigma_k$ .
15:         $\text{Grad}_{\mu_{3,k}} \leftarrow \text{Grad}_{\mu_{3,k}} - 2\epsilon_{ij}\beta_k\mu_{2,k}\sigma_k(1 - \sigma_k)\gamma_i$ .
16:         $\text{Grad}_{\mu_{4,k}} \leftarrow \text{Grad}_{\mu_{4,k}} - 2\epsilon_{ij}\beta_k\mu_{2,k}\sigma_k(1 - \sigma_k)$ .
17:         $\text{Grad}_{\mu_{5,k}} \leftarrow \text{Grad}_{\mu_{5,k}} - 2\epsilon_{ij}\eta_k\rho_j(\mu_{1,k} + \mu_{2,k}\sigma_k)$ .
18:         $\text{Grad}_{\mu_{6,k}} \leftarrow \text{Grad}_{\mu_{6,k}} - 2\epsilon_{ij}\rho_j(\mu_{1,k} + \mu_{2,k}\sigma_k)$ .
19:       end for
20:     end for
21:   end for
22:    $\text{Grad}_{\mu_0} \leftarrow \text{Grad}_{\mu_0} - 2\epsilon_{ij}$ .
23:   end for
24:   end for
25:   Update:  $\mu_0 \leftarrow \mu_0 - \alpha_0\text{Grad}_{\mu_0}$ 
26:   Update:  $\mu_{i,k} \leftarrow \mu_{i,k} - \alpha_i\text{Grad}_{\mu_{i,k}}, i = 1 : 6, k = 1 : N_c$ .
27: end for
28: Output: Final parameters  $\mu_i, i = 0 : 6$ .

```

models to optimize system parameters can impact system performance [27]. A more accurate fitting model typically leads to a more reliable and effective design.

By applying the same techniques as in [16], we introduce the unified loss model as in (4), where N_c is the number of terms. However, the MATLAB curve fitting tool does not support the multi-dimensional fitting with the structure given in (4). To find the parameters $\mu_0, \mu_{i,j}, i = 1 : 6, j = 1 : N_c$, we employ the gradient descent optimization method. The details of this approach are provided in Algorithm 1, where lines 4-5 to determine the fitting values ξ and the error between the fitted results ξ and the measured data in Tables I and IV, which is denoted as $\mathbf{Y}$. Lines 6 – 26 are to compute the gradients of ξ with respect to $\mu_i, i = 0 : 6$, while lines 27 and 28 are to update $\mu_i, i = 0 : 6$. In this paper, we set $\alpha_0 = \alpha_1 = \alpha_2 = 10^{-4}$, $\alpha_3 = \alpha_4 = 10^{-9}$, and $\alpha_5 = 10^{-10}$. The fitting model parameters corresponding to the results illustrated in Figs. 2, 3, and 4 are given in Tables XI, XII, and XIII (*placed at the end of the paper*), respectively. In addition, we examine the SemCom performance under Rician fading [20]; the corresponding empirical results and fitting results, are shown in Table XIV and Fig. 8 (*also placed at*

TABLE X: Average MSE between measured data and fitted values.

	Proposed	G. Sigmoid	Sum-Exp.
$\rho = 12$	0.0407	0.2474	0.0521
$\rho = 8$	0.0219	0.1425	0.1133

the end of the paper), respectively.

Figs. 2, 3, and 4 show that the proposed loss model fits well across all considered scenarios, including six different metrics (i.e., PSNR, MSE, SSIM, accuracy, precision, and recall) and two evaluation aspects (i.e., image reconstruction and image classification). Furthermore, we compare our approach with existing loss models to evaluate the fitting capability of the proposed unified loss model. In particular, the ‘G. Sigmoid’ corresponds to the generalized sigmoid model in [14], while the ‘Sum-Exp’ represents the sum-of-exponential model in [15]. These models are defined as

$$\xi_{\text{GSigmoid},\rho} = \kappa_{1,\rho} + \frac{\kappa_{2,\rho}}{1 + \exp(\kappa_{3,\rho}s + \kappa_{4,\rho})}, \quad (5)$$

$$\xi_{\text{SumExp},\rho} = \kappa_{5,\rho} \exp(\kappa_{6,\rho}s) + \kappa_{7,\rho} \exp(\kappa_{8,\rho}s) + \kappa_{9,\rho},$$

where $\kappa_{i,\rho}$ ’s are determined by fitting to empirical data corresponding to the CR ρ . It means both $\xi_{\text{GSigmoid},\rho}$ and $\xi_{\text{SumExp},\rho}$ are one-dimensional functions of the SNR γ , whereas our proposed model in (4) is a two-dimensional function that jointly depends on ρ and γ . Since the existing models capture only one-dimensional behavior with respect to SNR γ , we evaluate them by fixing $\rho = 8$ and $\rho = 12$, and then optimally determining the corresponding parameters $\kappa_{i,\rho}$ for each case. There is no restriction on determining $\kappa_{i,8}$ and $\kappa_{i,12}$. Fig. 5 and Table X illustrate the accuracy results obtained by applying DJSCC with EfficientViT. Our proposed model significantly outperforms both the ‘G. Sigmoid’ and ‘Sum-Exp’ models, even while capturing both dimensions simultaneously.

IV. A SEMANTIC LOSS MODEL-BASED RESOURCE ALLOCATION STUDY FOR EO MISSIONS

This section provides a resource allocation study that jointly optimizes transmission power and semantic compression to maximize the minimum user’s QoS in terms of image reconstruction quality. By leveraging our modeled loss models, the proposed approach enables a tractable solution for massive-MIMO Low Earth Orbit (LEO) satellite systems.

A. System Model

Consider a satellite-based imagery delivery network in which an imaging satellite directly transmits processed visual information to K single-antenna ground users (GUs) for real-time decision-making, such as disaster response and situational awareness. Let $\mathcal{K}$ denote the set of GUs, and B be the available bandwidth for the communications between the satellite and GUs. The satellite is equipped with onboard sensors to collect and capture the necessary data for fulfilling the monitoring tasks requested by the GUs. These tasks are considered in a

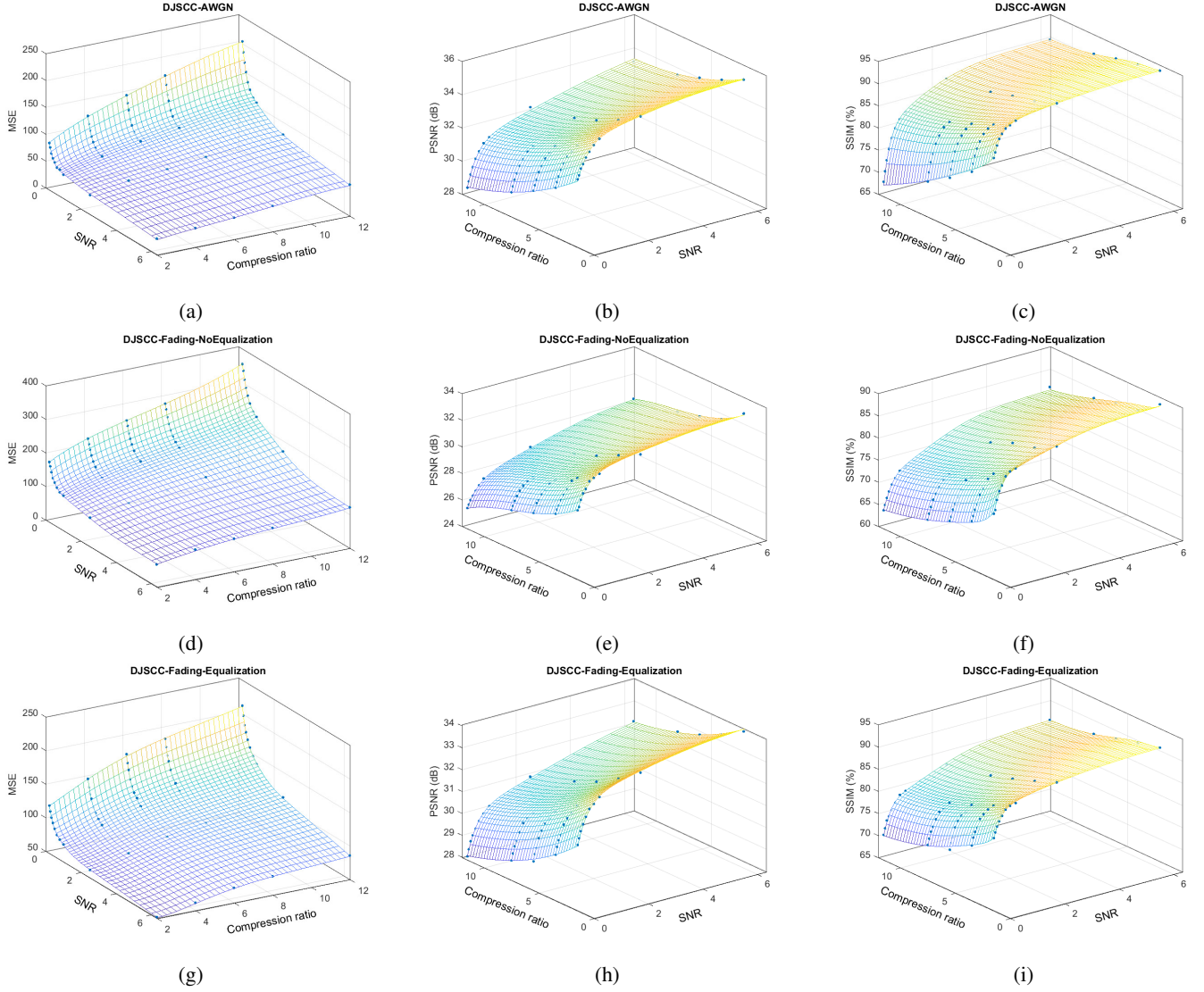


Fig. 2: Curve fitting models for the DJSCC case.

time window of T time-slots (TS)s, and let τ be the duration of a TS. In this SemCom system, we focus on the DJSCC for the image transmission. Particularly, the satellite employs the encoder to extract and encode semantic feature from captured images before transmitting them to the corresponding GUs. At the GU side, GU employs the decoder to decode received semantic features and reconstruct the corresponding images.

1) *Channel Model*: Assuming that the satellite is equipped with an M_a antenna array in a two-dimensional (2D) grid. Let $\mathbf{w}_{k,t} \in \mathbb{C}^{M_a \times 1}$ be the beamforming vector applied to user k in TS t , and it is normalized as $\|\mathbf{w}_{k,t}\|_2^2 = 1$. Furthermore, we adopt the sparse array implementation from [28] and the 2D non-uniform DFT (2D-NUDFT)-based beamforming and 2D-NUFFT techniques from [29] to improve beam alignment and reduce sidelobe gains. Assume that at TS t , the center direction of l -th beam is $\eta_{t,l}$. Let δ_t^S and $\delta_{k,t,U}$ denote the locations of the satellite and GU k at TS t , respectively. Then, the boresight angle of n -th beam from the viewpoint of the

satellite to GU k at TS t is given by

$$\omega_{k,l,t} = \arccos \left(\frac{(\delta_t^S - \eta_{t,l})^\top (\delta_t^S - \delta_{k,t,U})}{\|\delta_t^S - \eta_{t,l}\| \|\delta_t^S - \delta_{k,t,U}\|} \right). \quad (6)$$

It can be verified that when $\eta_{t,l} = \delta_{k,t,U}$, the boresight angle $\omega_{k,l,t} = 0$. The channel between the satellite and GU k is modeled as in (7), where $PL_{k,t} = 1/(4\pi f_c d_{k,t}/\nu_c)^2$ represents the free-space path loss, and $d_{k,t}$ denotes the distance between the satellite and GU k at TS t . The term $\nu_{k,t}$ denotes the Doppler shift of GU k due to the satellite's mobility at TS t . The parameters G^{GU} and G^{SAT} are the antenna gains of the GU and the satellite, respectively. The function $f_{\text{bsa}}(\omega_{k,l,t})$, reflecting the boresight angle impact, is expressed as in [30]:

$$f_{\text{bsa}}(\omega) = \begin{cases} 1, & \text{if } \omega = 0, \\ 4 \left| \frac{\mathcal{J}_1(b_{\text{ap}} a_{\text{ap}} \sin(\omega))}{b_{\text{ap}} a_{\text{ap}} \sin(\omega)} \right|^2, & \text{if } \omega \neq 0, \end{cases}$$

where a_{ap} is the antenna aperture, $\mathcal{J}_1(\cdot)$ is the first-order Bessel function of the first kind, and $b_{\text{ap}} = 2\pi f_c/c$. The terms $\varrho_{k,t}^c$ and ϱ_t^r represent attenuation due to cloud and rain, respectively, which are modeled according to [31].

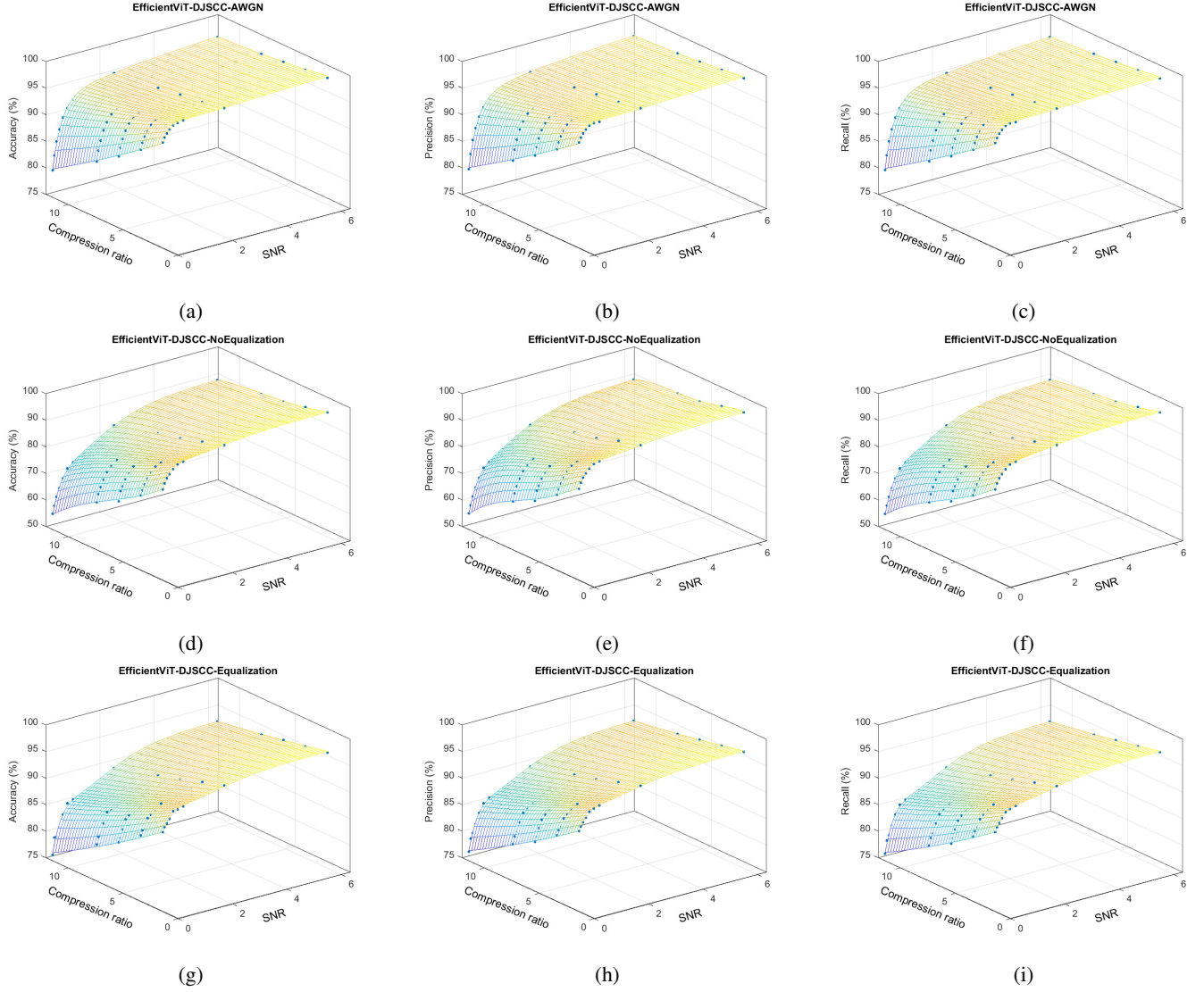


Fig. 3: Curve fitting models for the EfficientViT-DJSCC case.

2) *DJSCC-based SemCom Model*: Let $\mathbf{x}_k$ denote the original source image data that the satellite senses and collects for the EO application of GU k during the considered interval. At the satellite, with input $\mathbf{x}_k$, the encoder extracts and encodes the semantic feature into the symbols $\mathbf{z}_k$. Let $\zeta_{k,i}$ be the number of source data symbols $\mathbf{x}_k$, and $\zeta_{k,o}$ be the number of transmitted symbols $\mathbf{z}_k$. Particularly, let $E(\cdot|\rho)$ be the encoder function, where $\rho = \zeta_{k,i}/\zeta_{k,o}$ is the CR parameter. The output symbol for GU k is given by $\mathbf{z}_k = E(\mathbf{x}_k|\rho_k), \forall k \in \mathcal{K}$. For convenience, let's denote $\boldsymbol{\rho} = [\rho_k]_{\forall k \in \mathcal{K}}$. It is assumed that the CRn is bounded by the following condition:

$$(C1): \quad \rho_k \in [\rho^{\min}, \rho^{\max}], \quad (8)$$

where $\rho^{\min}$ and $\rho^{\max}$ denote the minimum and maximum allowable CR. Subsequently, $\mathbf{z}_k$ is prepared to be transmitted over wireless environment to GU k . Let $\bar{\mathbf{z}}_{k,t}$ be the transmitted symbol for GU k at TS t which is constructed from $\mathbf{z}_k$ and normalized as $\mathbb{E}[\|\bar{\mathbf{z}}_{k,t}\|^2] = 1$. The received symbol at GU k at TS t is expressed as

$$\bar{\mathbf{y}}_{k,t} = \sqrt{p_{k,t}} \mathbf{w}_{k,t}^H \mathbf{h}_{k,t} \bar{\mathbf{z}}_{k,t} + \mathbf{n}_{k,t}, \quad (9)$$

where $p_{k,t}$ and $\mathbf{h}_{k,t}$ are the transmit power and channel coefficients from LSAT to GU k at TS t , and $\mathbf{n}_{k,t} \sim \mathcal{CN}(0, \sigma_k^2 \mathbf{I})$ is the AWGN. The corresponding SINR is expressed as

$$\gamma_{k,t} = \frac{p_{k,t} |\mathbf{w}_{k,t}^H \mathbf{h}_{k,t}|^2}{\sum_{j \neq k} p_{j,t} |\mathbf{w}_{j,t}^H \mathbf{h}_{k,t}|^2 + \sigma_k^2}. \quad (10)$$

Additionally, regarding limited power budget, denoted by $p_{\max}$, the transmit power must satisfy

$$(C2): \quad \sum_{\forall k \in \mathcal{K}} p_{k,t} \leq p_{\max}, \quad \forall t.$$

The constraint indicates that $\zeta_{k,o}$ is transmitted within T TSs is expressed as follows

$$(C3): \quad \zeta_{k,o} \leq \sum_{\forall t} B\tau \|p_{k,t}\|_0,$$

where $\|x\|_0$ indicates ℓ_0 -norm.

Once the transmission of $\mathbf{z}_k$ is completed, the received symbols $\bar{\mathbf{y}}_{k,t}$ corresponding to $\mathbf{z}_k$ are aggregated into the symbol vector $\mathbf{y}_k$. Subsequently, GU k applies a decoder to reconstruct the corresponding image. The semantic indication is approximated by the loss model (4).

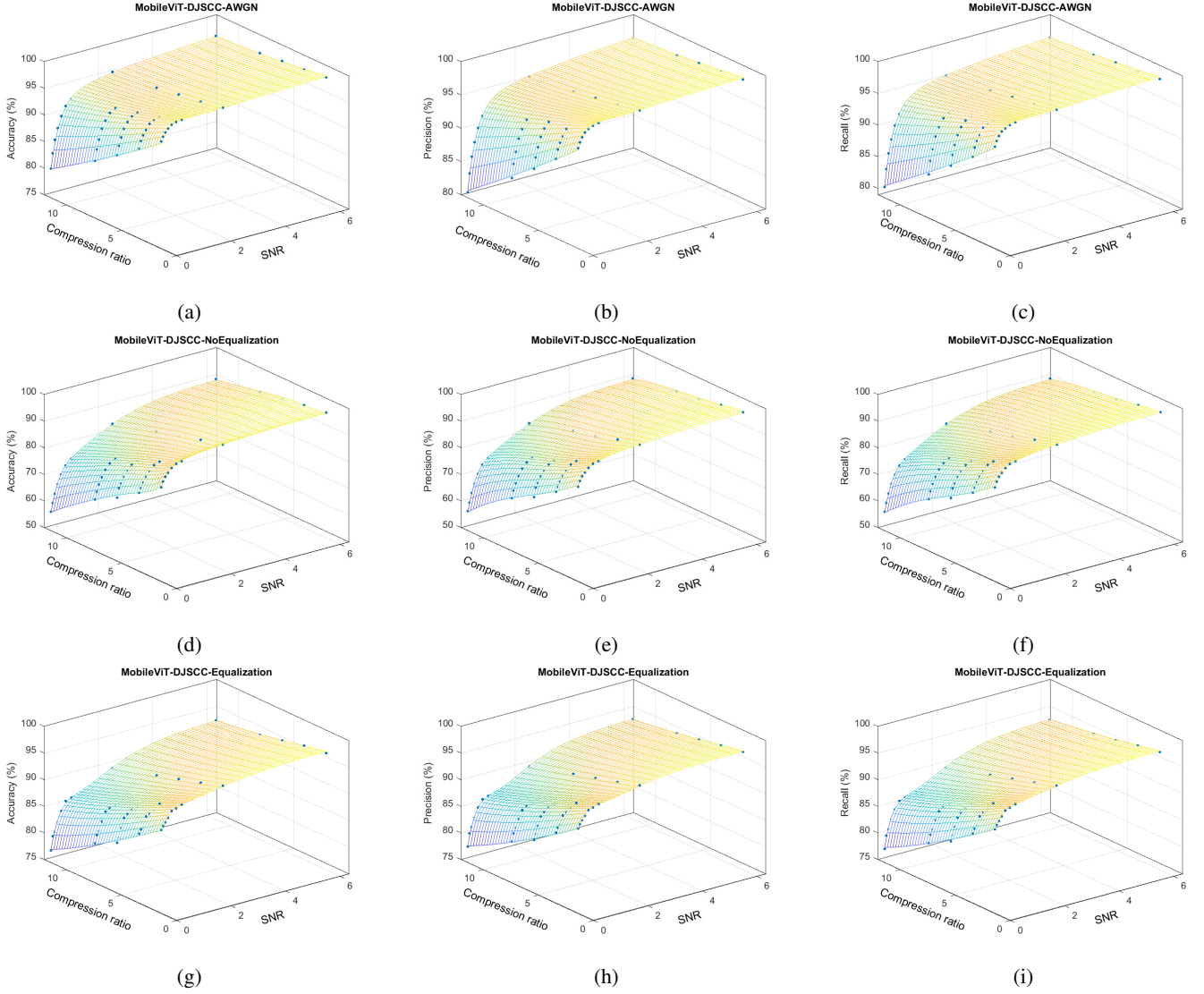


Fig. 4: Curve fitting models for the MobileViT-DJSCC case.

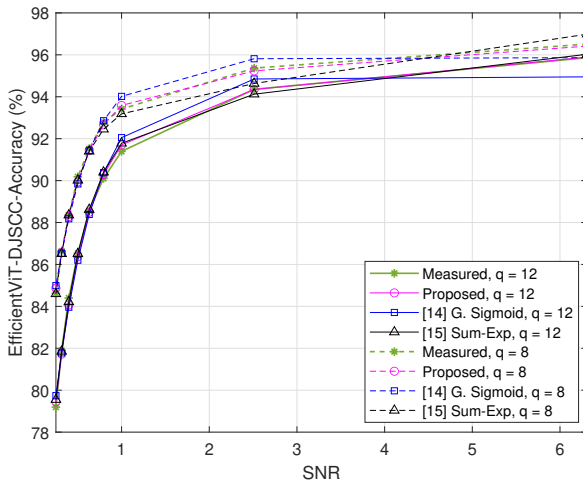


Fig. 5: A comparison of different fitting models.

Let $\bar{\gamma}_k$ be the lower-bound of SINR of GU k in the

considered interval. Specifically, it is defined as follows

$$\bar{\gamma}_k = \min\{\gamma_{k,t} | p_{k,t} > 0\}. \quad (11)$$

Then, $\xi(\rho_k, \bar{\gamma}_k)$ reflects the worst performance of GU k .

3) *Problem Formulation*: In this model, we aim to maximize the minimum worst weighted semantic performance of GUs by optimizing the CR and the transmit power. Let $\mathbf{p} = \{p_{k,t}\}_{\forall k,t}$, $\boldsymbol{\rho} = \{\rho_k\}_{\forall k}$, and λ_k be the weight of user k , the optimization problem is formulated as:

$$(\mathcal{P}) \max_{\mathbf{p}, \boldsymbol{\rho}} \min_k u \lambda_k \xi(\rho_k, \bar{\gamma}_k) \text{ s.t. } (C1) - (C3),$$

where $u = -1$ for MSE metric considered, otherwise $u = 1$.

B. Algorithm Development

1) *Problem Transformation*: To gain insight into its non-smooth min-max objective function, we recast $(\mathcal{P})$ into the

$$\mathbf{h}_{k,l,t} = \frac{\sqrt{G^{\text{GU}} G^{\text{SAT}} f_{\text{bsa}}(\omega_{k,l,t}) \text{PL}_{k,t}}}{\varrho_{k,t}^c \varrho_t^f} \exp \left[j2\pi \left(t\tau\nu_{k,t} - \frac{f_c d_{k,t}}{\nu_c} \right) \right] e^{j2\pi(t\tau\nu_k^t - f_c d_k(t)/\nu_c)} \mathbf{a}(\theta_{k,t}, \phi_{k,t}), \quad (7)$$

following equivalent problem:

$$(\mathcal{P}_1) \quad \max_{\mathbf{p}, \rho, \psi, \bar{\gamma}} \quad \psi \quad \text{s.t.} \quad (C1) - (C3),$$

$$(C4) : \quad \|p_{k,t}\|_0 \bar{\gamma}_k \leq \gamma_{k,t}, \forall k, t,$$

$$(C5) : \quad \psi \leq \lambda_k \xi(\rho_k, \bar{\gamma}_k), \forall k,$$

where $\bar{\gamma} = \{\bar{\gamma}_k, \forall k\}$. Problem $(\mathcal{P}_1)$ is still difficult to solve due to the nonconvex constraint $(C4)$ and $(C5)$ and the sparse ℓ_0 -norm terms. To address these challenges, an iterative solution has been proposed as in the following sub-section.

2) *Many-to-Many Matching with Externalities*: To address the ℓ_0 -norm components in $(C3)$ and $(C4)$, a many-to-many matching algorithm between the users and selected TSs is developed. Let $\mathcal{T} = \{1, \dots, T\}$. Each user $k \in \mathcal{K}$ will select a set of TSs $\mathcal{T}_k \in \mathcal{T}$ for the data transmission. The set of selected users for the transmission in TS t is denoted as $\mathcal{K}_t = \{k \in \mathcal{K} | t \in \mathcal{T}_k\}$. Due to equations (4) and (10), the preference value in the matching for user k not only depends on its selection that it is matched with but also on the matching of other $k' \in \mathcal{K}$. According to [32], the matching game with externalities can be used to address this problem.

Specifically, let μ be a two-direction matching between two sets $\mathcal{K}$ and $\mathcal{T}$, i.e. $\mu(k)|_{k \in \mathcal{K}} \subset \mathcal{T}$ and $\mu(t)|_{t \in \mathcal{T}} \subset \mathcal{K}$. In addition, one has: i) $|\mu(k)| \leq T, \forall k \in \mathcal{K}$, and $|\mu(k)| = \emptyset$ if user k is not scheduled for its transmission in any TSs; ii) $|\mu(t)| \leq K, \forall t \in \mathcal{T}$, and $|\mu(t)| = \emptyset$ if TS t is not scheduled for the transmission of any users; iii) $t \in \mu(k)$ iff $k \in \mu(t)$. Let $\Omega = \{(k, t)\}_\mu$ be the set of all matching links between $\mathcal{K}$ and $\mathcal{T}$ following the matching function μ .

Definition 1. We further define a concept of swap matching, $\mu_{(k,t)}^{(k',t')}$ whose set of all matching links is given as follows:

$$\Omega_{\mu_{(k,t)}^{(k',t')}} = \{\Omega_\mu \setminus \{(k, t), (k', t')\}\} \cup \{(k, t'), (k', t)\}, \quad (14)$$

where $t = \mu(k)$ and $t' = \mu(k')$. Noted that the swap matching $\mu_{(k,t)}^{(k',t')}$ also requires the swap of the transmit power.

To develop the matching game with externalities, we first define the preference value $U(\mu)$ of the matching μ as $U(\mu) = \min_{j \in \mathcal{K}} \lambda_k \xi(\rho_j, \bar{\gamma}_j)$, where $\rho_j = \frac{\zeta_{j,j}}{B\tau|\mu(j)|}$ and $\bar{\gamma}_j = \min\{\gamma_{j,t} | t \in \mu(j)\}$. Specifically, if $p_{k',t'} \neq 0$ and $p_{k,t} \neq 0$, we swap $p_{k,t'} = p_{k',t'}$, $p_{k',t} = p_{k,t}$ and keep remaining transmit powers. If $p_{k',t'} = 0$, $t' \neq \emptyset$, and $k \neq \emptyset$, we assign $p_{k,t'} = \frac{\sum_j p_{j,t'} |\mu|}{|\mu(t')|+1}$ and $p_{j,t'} = p_{j,t} \frac{|\mu(t')|}{|\mu(t')|+1}, \forall j \in \mu(t')$. If $p_{k,t} = 0$, $t \neq \emptyset$, and $k' \neq \emptyset$, we assign $p_{k',t} = \frac{\sum_j p_{j,t} |\mu|}{|\mu(t)|+1}$ and $p_{j,t} = p_{j,t} \frac{|\mu(t)|}{|\mu(t)|+1}, \forall j \in \mu(t)$. Then, we propose Algorithm 2 to determine the matching μ , where the swap-blocking pair procedure in Definition 2 [32] is adopted.

Definition 2. Let $k \in \mathcal{K}$ and $k' \in \mathcal{K} \cup \mathcal{O}$, (k, k') be a swap-blocking pair if and only if 1) $\forall x \in \{(k, t)|_{k \in \mathcal{K}, t = \mu(k)}, (k', t')|_{k' \in \mathcal{K}, t' = \mu(k')}\}$, $U(\mu_{k'}^{k'}) \leq U(\mu)$,

Algorithm 2 Time-slot Matching Algorithm

```

1: Initialize: Variables  $\|p_{k,t}\|_0$  for all  $k, t$ .
2: repeat
3:   For each  $k$ , search for another  $k'$  or an open spot  $\mathcal{O}$  to form a swap-blocking pair.
4:   if  $\{(k, t), (k', t')\}, \{(k, t), (\emptyset, \emptyset)\}$ , or  $\{(k, \emptyset), (\emptyset, t')\}$  forms a swap-blocking pair then
5:     Update  $\mu \leftarrow \mu_{(k,t)}^{(k',t')}$ ,  $\mu \leftarrow \mu_{(k,t)}^{(\emptyset, \emptyset)}$ , or  $\mu \leftarrow \mu_{(k,\emptyset)}^{(\emptyset, t')}$ , respectively.
6:   end if
7: until convergence

```

Algorithm 3 Power Allocation Algorithm

```

1: Initialize:  $\varrho, \psi^{\max}$  and  $\psi^{\min}$ .
2: repeat
3:   Assign  $\psi = \frac{\psi^{\max} + \psi^{\min}}{2}$ .
4:   Determine  $\bar{\gamma}_k$  such that  $\lambda_k \xi(\rho_k, \bar{\gamma}_k) = \psi$  where  $\rho_k = \frac{\zeta_{k,i}}{B\tau|\mu(k)|}$ .
5:   Solve problem  $(\mathcal{P}_2)$  to find  $\epsilon$ .
6:   if  $\epsilon < 0$  then Assign  $\psi^{\max} = \psi$  else Assign  $\psi^{\min} = \psi$  endif.
7: until  $\psi^{\max} - \psi^{\min} \leq \varrho$ 

```

Algorithm 4 Resource Allocation Algorithm in LEO Systems with DJSCC

```

1: Initialize:  $p_{k,t}$  for all  $k, t$ .
2: repeat
3:   Determine matching  $\mu$  using Algorithm 2.
4:   Determine the transmit power using Algorithm 3.
5: until convergence

```

and 2) $\exists(k, k^*)$ in such a set that equality cannot occur, e.g., $U(\mu_{k^*}^{k^*}) < U(\mu)$.

3) *Bisection Search*: The objective of problem $(\mathcal{P}_1)$ is a monotonic functions of ψ . Therefore, we employ the bisection search to find the optimum ψ^* of $(\mathcal{P}_1)$ where upper bound $\psi^{\max}$ and lower bound $\psi^{\min}$ are iteratively updated until the difference between them becomes sufficiently small. These bounds are updated according to the feasibility verification, as shown in Algorithm 3. Specifically, for a given ψ and matching μ and due to the monotonic property of $\xi(\rho_k, \bar{\gamma}_k)$, we determine $\bar{\gamma}_k$ that satisfies $\lambda_k \xi(\rho_k, \bar{\gamma}_k) = \psi$. Then, constraint $(C4)$ is rewritten as follows

$$(\tilde{C}4) : p_{k,t} |\mathbf{w}_{k,t}^H \mathbf{h}_{k,t}|^2 - \bar{\gamma}_k \left(\sum_{j \neq k} p_{j,t} |\mathbf{w}_{j,t}^H \mathbf{h}_{k,t}|^2 + \sigma_k^2 \right) \geq 0.$$

To perform feasibility for constraints $(C2)$ and $(\tilde{C}4)$, we consider the problem: $(\mathcal{P}_2) \max_{\mathbf{p}, \epsilon} \epsilon$ s.t. $(C2)$, $(\tilde{C}4)$, where

$$(\tilde{C}4) : p_{k,t} |\mathbf{w}_{k,t}^H \mathbf{h}_{k,t}|^2 - \bar{\gamma}_k \left(\sum_{j \neq k} p_{j,t} |\mathbf{w}_{j,t}^H \mathbf{h}_{k,t}|^2 + \sigma_k^2 \right) \geq \epsilon.$$

It can be verified that when $\epsilon \geq 0$, constraint $(\tilde{C}4)$ is not satisfied. Accordingly, the upper bound $\psi^{\max}$ and lower bound $\psi^{\min}$ will be updated as in Algorithm 3. Furthermore, problem $(\mathcal{P}_2)$ is a standard linear programming (LP) problem, which can be solved efficiently by using CVX-solvers.

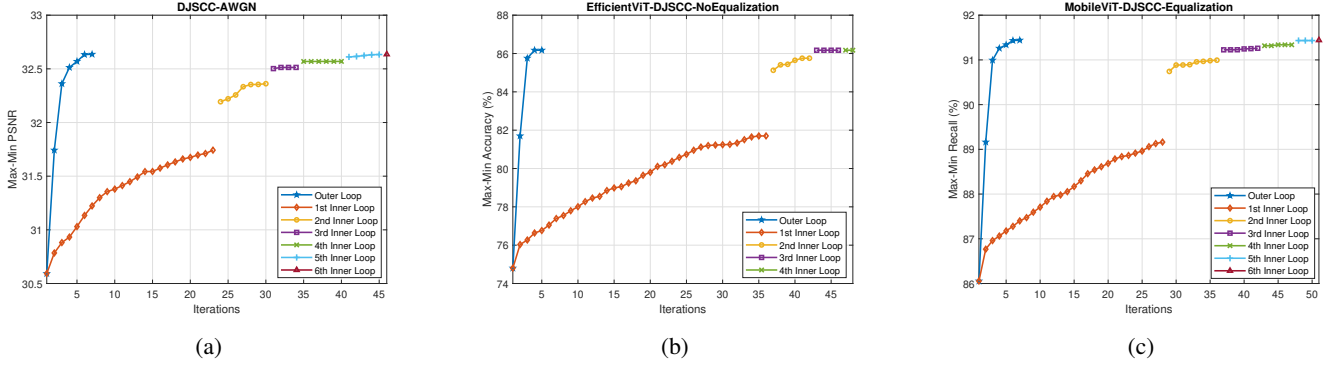


Fig. 6: Convergence of the Proposed Algorithms.

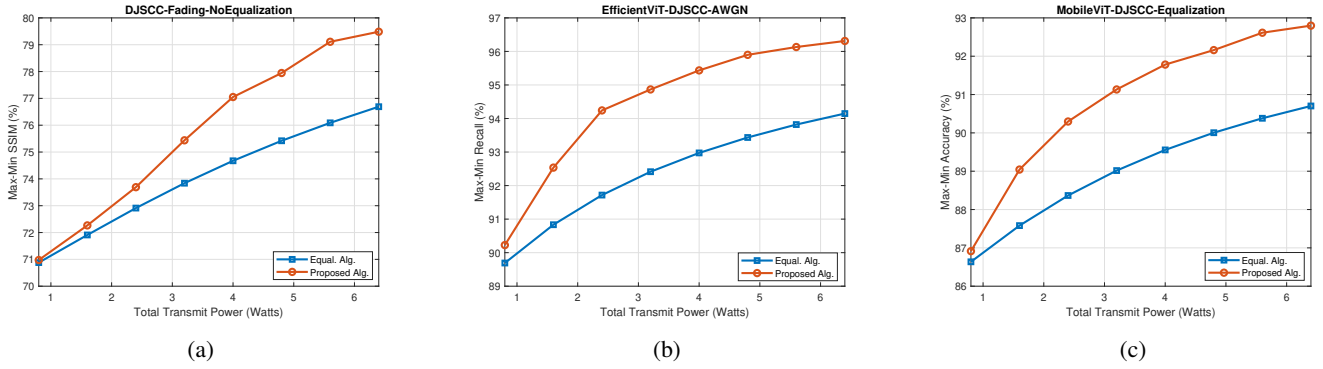


Fig. 7: Performance w.r.t. Maximum Transmit Power.

C. Numerical Results

In this section, the numerical results are studied to evaluate the proposed algorithms in various scenarios. The EUROSAT dataset [33] is used for the LEO satellite EO simulation. The simulated system is examined in a time-window of 100 seconds. The frequency $f_c = 18$ GHz, bandwidth $B = 20$ MHz, TS duration $\tau = 1$ second, satellite altitude at 500 km, satellite antenna gain $G^{\text{SAT}} = 37.1$ dBi, GU antenna gain, $G^{\text{GU}} = 0$ dBi, number of GUs and TSs, $(K, T) = (20, 50)$, maximum transmit power $p_{\max} = 4$ Watts, number of antennas $M_a = 34$, and user weight $\lambda_k = 1$. EUROSAT dataset [33] is considered for experiment results. The cloud water content $\rho_{k,\text{water}}$ follows a uniform random distribution between 0.1 and 0.15. Users are distributed randomly within a circular cell with a radius of 100 km. To evaluate the performance of the proposed algorithm, we also compare it with a benchmark scheme, referred to as the 'Equal. Alg.', in which all users are served in every TS and the transmit power is equally allocated among all users. Note that this benchmark 'Equal Alg.' is also used to generate the initial point for Algorithm 4.

Fig. 6 illustrates the convergence behavior of the proposed algorithm. Three scenarios are randomly selected from the 27 considered scenarios to demonstrate convergence: namely, 'DJSCC-AWGN' with the performance metric 'PSNR', 'EfficientViT-DJSCC-NoEqualization' with 'Accuracy', and 'MobileViT-DJSCC-Equalization' with 'Recall'. The outer loop depicts the max-min values of the corresponding metric across iterations of Algorithm 4, while the inner loops show the max-min values across iterations of Algorithm 2.

The number of iterations of Algorithm 3 is implicitly reflected by the gaps between consecutive inner loops, where ϱ is set to 10^{-1} . The source data volume is set equal to $4BT\tau$. The figure shows that the proposed Algorithm 4 converges within five outer-loop iterations. Moreover, the iteration numbers of inner loops decrease as the outer-loop iterations progress; eventually, it is less than three in the fourth inner loop.

Fig. 7 shows a comparison between the proposed algorithm (i.e., Algorithm 4) and the benchmark 'Equal. Alg.'. As the maximum transmit power increases, both algorithms achieve improved semantic performance; however, the proposed algorithm consistently outperforms the benchmark across the entire power range. This performance gain highlights the effectiveness of the proposed power allocation and scheduling strategy in enhancing overall user fairness and QoS.

V. CONCLUSIONS

This work establishes a principled semantic loss modeling framework for DJSCC-based imagery transmission. By learning the implicit mapping between JSCC CR, channel signal-to-noise ratio, and semantic performance, the proposed data-driven model provides a unified and tractable characterization of semantic loss that encompasses both reconstruction-centric quality degradation and task-oriented performance objectives. Unlike conventional distortion metrics or task-specific heuristics, the proposed semantic loss model is explicitly designed to be optimization-friendly, enabling its direct integration into cross-layer resource allocation and system-level decision-making. Building upon this foundation, a semantic-aware

resource allocation framework is developed to maximize the minimum weighted worst-case semantic performance across ground users, demonstrating how semantic loss modeling can be systematically exploited to guide power and bandwidth allocation under heterogeneous channel and application requirements. Extensive numerical results confirm that the proposed unified model achieves superior fitting accuracy and robustness compared to existing loss formulations, while enabling consistent performance gains when embedded into satellite resource management strategies. Overall, this work bridges the gap between SemCom modeling and RMS optimization, positioning semantic loss as a first-class design metric for future imagery transmission-centric wireless systems operating under stringent bandwidth, power, and link variability constraints.

REFERENCES

- [1] I. Leyva-Mayorga, M. Martínez-Gost, M. Moretti, A. Pérez-Neira, M. Á. Vázquez, P. Popovski, and B. Soret, "Satellite edge computing for real-time and very-high resolution earth observation," *IEEE Trans. Commun.*, vol. 71, no. 10, pp. 6180–6194, 2023.
- [2] G. Fontanesi *et al.*, "Artificial intelligence for satellite communication and non-terrestrial networks: A survey," *arXiv preprint arXiv:2304.13008*, 2023.
- [3] M. Affek and J. Szymański, "A survey on the datasets and algorithms for satellite data applications," *IEEE J. Sel. Topics Applied Earth Obs. Remote Sensing*, 2024.
- [4] H.-f. Chou, S. Solanki, V. N. Ha, L. Chen, S. L. Ma, H. Al-Hraishawi, G. Eappen, and S. Chatzinotas, "Edge ai empowered physical layer security for 6g ntn: Potential threats and future opportunities," *arXiv preprint arXiv:2401.01005*, 2023.
- [5] W. Yang, H. Du, Z. Q. Liew, W. Y. B. Lim, Z. Xiong, D. Niyato, X. Chi, X. Shen, and C. Miao, "Semantic communications for future internet: Fundamentals, applications, and challenges," *IEEE Commun. Surveys & Tuts.*, vol. 25, no. 1, pp. 213–250, 2022.
- [6] H.-f. Chou *et al.*, "On-air deep learning integrated semantic inference models for enhanced earth observation satellite networks," *arXiv preprint arXiv:2409.15246*, 2024.
- [7] N. Li, Y. Deng, and D. Niyato, "Goal-oriented semantic communication for wireless video transmission via generative AI," *IEEE Trans. Wireless Commun.*, 2026.
- [8] L. Guo, W. Chen, Y. Sun, B. Ai, N. Pappas, and T. Quek, "Diffusion-driven semantic communication for generative models with bandwidth constraints," *IEEE Trans. Wireless Commun.*, 2025.
- [9] E. Boursoulatz, D. B. Kurka, and D. Gündüz, "Deep joint source-channel coding for wireless image transmission," *IEEE Trans. Cog. Commun. Netw.*, vol. 5, no. 3, pp. 567–579, 2019.
- [10] C. Bian, Y. Shao, and D. Gündüz, "A deep joint source-channel coding scheme for hybrid mobile multi-hop networks," *IEEE J. Sel. Areas Commun.*, 2025.
- [11] J. Xu, B. Ai, N. Wang, and W. Chen, "Deep joint source-channel coding for csi feedback: An end-to-end approach," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 1, pp. 260–273, 2022.
- [12] J. Dai, S. Wang, K. Tan, Z. Si, X. Qin, K. Niu, and P. Zhang, "Nonlinear transform source-channel coding for semantic communications," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 8, pp. 2300–2316, 2022.
- [13] J. Huang, K. Yuan, C. Huang, and K. Huang, " D^2 -JSCC: Digital deep joint source-channel coding for semantic communications," *IEEE J. Sel. Areas Commun.*, 2025.
- [14] W.-F. Wang, J.-P. Sheu, and N. Van Cuong, "Joint scheduling and relay allocation for semantic and delay-sensitive wireless communications," *IEEE Trans. Green Commun. Netw.*, vol. 10, pp. 1666–1678, 2025.
- [15] G. Ding, S. Liu, J. Yuan, and G. Yu, "Joint URLLC traffic scheduling and resource allocation for semantic communication systems," *IEEE Trans. Wireless Commun.*, vol. 23, no. 7, pp. 7278–7290, 2023.
- [16] T. T. Nguyen *et al.*, "A semantic-loss function modeling framework with task-oriented machine learning perspectives," *Proc. IEEE Int. Conf. Machine Learn. Commun. Netw. (ICMLCN)*, 2025.
- [17] P. Helber, B. Bischke, A. Dengel, and D. Borth, "Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification," *IEEE J. Sel. Top. Appl. Earth. Obs. Remote Sens.*, vol. 12, no. 7, pp. 2217–2226, 2019.
- [18] X. Liu, H. Peng, N. Zheng, Y. Yang, H. Hu, and Y. Yuan, "Efficientvit: Memory efficient vision transformer with cascaded group attention," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 14 420–14 430, 2023.
- [19] S. Mehta and M. Rastegari, "Separable self-attention for mobile vision transformers," *Trans. Machine Learn. Research*, 2023.
- [20] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.
- [21] L. X. Nguyen, S. S. Hassan, Y. K. Tun, K. Kim, Z. Han, and C. S. Hong, "Semantic communication enabled 6G-NTN framework: A novel denoising and gateway hop integration mechanism," *IEEE Trans. Wireless Commun.*, 2025.
- [22] M. Zhang, G. Zhu, R. Jin, X. Chen, Q. Shi, C. Zhong, and K. Huang, "Beamforming design for semantic-bit coexisting communication system," *IEEE J. Sel. Areas Commun.*, 2025.
- [23] H. Hu, K. Song, R. Fan, C. Zhan, J. Xu, and J. Yang, "Energy-efficient image semantic communication: Architecture design and optimal joint allocation of communication and computation resources," *IEEE Trans. Circuits Syst. Video Technol.*, 2025.
- [24] M. Karaa, H. Ghazzai, G. Setti, and L. Sboui, "Efficient resource allocation for semantic video surveillance transmission over LEO satellites," in *IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. (PIMRC)*. IEEE, 2025, pp. 1–6.
- [25] S. S. Hassan, L. X. Nguyen, Y. K. Tun, Z. Han, and C. S. Hong, "Semantic enabled 6G leo satellite communication for earth observation: A resource-constrained network optimization," in *IEEE Global Commun. Conf. (GLOBECOM)*. IEEE, 2024, pp. 4472–4478.
- [26] T. D. Le *et al.*, "On-board satellite image classification for earth observation: A comparative study of vit models," *arXiv preprint arXiv:2409.03901*, 2024.
- [27] T. T. Nguyen, V. N. Ha, L. B. Le, and R. Schober, "Joint data compression and computation offloading in hierarchical fog-cloud systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 293–309, 2019.
- [28] J. A. Vázquez-Peralvo, J. C. M. Duncan, V. N. Ha, R. Palisetty, and S. Chatzinotas, "Planar sparse arrays using triangular lattice transformations," in *IEEE Int. Symp. Antennas Propag. (ISAP)*. IEEE, 2024.
- [29] A. Capozzoli, C. Curcio, and A. Liseno, "Optimized nonuniform FFTs and their application to array factor computation," *IEEE Trans. Antenna Propag.*, vol. 67, no. 6, pp. 3924–3938, 2018.
- [30] "Study on new radio (nr) to support non-terrestrial networks," 3rd Generation Partnership Project (3GPP), Technical Report TR 38.811, Sep. 2020.
- [31] H. Nguyen-Kha, V. N. Ha, E. Lagunas, S. Chatzinotas, and J. Grotz, "Joint two-tier user association and resource management for integrated satellite-terrestrial networks," *IEEE Trans. Wireless Commun.*, 2024.
- [32] J. Zhao, Y. Liu, K. K. Chai, Y. Chen, and M. El-kashlan, "Many-to-many matching with externalities for device-to-device communications," *IEEE Wireless Commun. Letters*, vol. 6, no. 1, pp. 138–141, 2016.
- [33] P. Helber, B. Bischke, A. Dengel, and D. Borth, "Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification," *IEEE J. Sel. Top. Appl. Earth. Obs. Remote Sens.*, 2019.

TABLE XI: PARAMETERS OF THE PROPOSED MODEL IN FIG. 2.

$i \backslash$ Parameters	μ_0	$\mu_{i,1}$	$\mu_{i,2}$	$\mu_{i,3}$	$\mu_{i,4}$	$\mu_{i,5} \times 10^3$	$\mu_{i,6}$
Model in Fig. 2 under DJSCC-AWGN conditions							
Image-MSE							
1	-132.851	325.366	-308.162	-2.900	-2.386	-267.139	0.000
2		-206.275	549.240	2.640	-0.632	61.433	0.000
3		-284.726	-194.664	-1.172	-0.607	-328.019	0.000
4		-161.191	204.884	-0.080	1.757	38.637	0.000
5		-65.899	61.225	2.840	5.172	222.567	0.000
6		929.199	-1160.623	2.247	0.330	57.890	0.000
Image-PSNR							
1	29.574	36.649	-3.168	0.632	-3.436	-39.516	0.000
2		-104.058	-4.692	0.156	-13.865	-12.833	0.000
3		-11.463	65.229	2.657	3.110	22.610	0.000
4		-55.011	-4.879	-0.433	1.825	19.184	0.000
5		-26.117	28.747	-0.269	20.030	-7.904	0.000
6		75.898	-14.652	-2.346	-2.705	-7.300	0.000
Image-SSIM							
1	28.768	126.256	-15.514	-1.740	19.184	-308.684	0.000
2		55.943	-11.283	-0.996	0.955	19.585	0.000
3		153.574	4.197	-0.213	-7.228	-701.928	0.000
4		-206.346	16.021	-1.375	1.974	-548.854	0.000
5		-37.987	6.056	0.664	5.817	-339.823	0.000
6		-33.014	28.336	4.731	0.568	78.149	0.000
Model in Fig. 2 under DJSCC-Fading-NoEqualization conditions							
Image-MSE							
1	200.439	-407.191	-406.506	-1.454	-1.663	-443.539	0.000
2		485.154	-35.573	1.184	-2.577	97.516	0.000
3		-155.989	-445.938	3.326	1.416	83.938	0.000
4		-159.044	564.760	1.519	1.253	-476.378	0.000
5		1338.297	-1522.123	2.072	1.836	-132.172	0.000
6		125.958	66.464	3.080	-1.121	-16.486	0.000
Image-PSNR							
1	31.544	40.347	-1.528	0.631	-3.384	38.903	0.000
2		-104.629	-4.692	0.156	-13.865	-5.626	0.000
3		-10.415	42.200	2.669	2.593	-91.949	0.000
4		-46.070	-3.120	-0.482	1.713	64.068	0.000
5		-25.334	29.530	-0.269	20.030	1.612	0.000
6		78.874	-21.075	-2.350	-2.555	34.630	0.000
Image-SSIM							
1	27.109	122.020	-19.739	-1.740	20.383	-317.440	0.000
2		56.389	-11.942	-0.996	1.748	-1.479	0.000
3		160.652	4.184	-0.213	-6.995	-699.455	0.000
4		-198.558	13.611	-1.375	2.289	-545.566	0.000
5		-40.230	4.013	0.664	7.146	-337.370	0.000
6		-22.845	18.972	4.731	-0.692	-10.006	0.000
Model in Fig. 2 under DJSCC-Fading-Equalization conditions							
Image-MSE							
1	320.647	107.294	-313.107	-2.898	-2.409	-633.215	0.000
2		23.089	503.126	2.738	-0.055	53.990	0.000
3		-838.890	-108.233	-1.172	-0.476	-439.553	0.000
4		-116.662	228.270	-0.072	1.754	76.759	0.000
5		291.115	416.005	2.836	5.135	-490.789	0.000
6		723.096	-1520.359	2.314	1.120	46.718	0.000
Image-PSNR							
1	29.612	36.433	-2.594	0.633	-3.435	-57.122	0.000
2		-104.363	-4.692	0.156	-13.865	-22.813	0.000
3		-10.742	63.188	2.658	3.112	0.314	0.000
4		-54.300	-4.551	-0.430	1.829	6.866	0.000
5		-26.273	28.591	-0.269	20.030	-7.407	0.000
6		75.753	-12.036	-2.346	-2.706	-6.688	0.000
Image-SSIM							
1	27.109	122.020	-19.739	-1.740	20.383	-317.440	0.000
2		56.389	-11.942	-0.996	1.748	-1.479	0.000
3		160.652	4.184	-0.213	-6.995	-699.455	0.000
4		-198.558	13.611	-1.375	2.289	-545.566	0.000
5		-40.230	4.013	0.664	7.146	-337.370	0.000
6		-22.845	18.972	4.731	-0.692	-10.006	0.000

TABLE XII: PARAMETERS OF THE PROPOSED MODEL IN FIG. 3.

$i \backslash$ Parameters	μ_0	$\mu_{i,1}$	$\mu_{i,2}$	$\mu_{i,3}$	$\mu_{i,4}$	$\mu_{i,5} \times 10^3$	$\mu_{i,6}$
Model in Fig. 3 under EfficientViT-DJSCC-AWGN conditions							
EfficientViT-Accuracy							
1	62.683	-128.798	-7.955	0.859	2.765	-1.475	0.008
2		-16.487	63.019	0.253	3.622	-11.792	0.166
3		-124.914	4.798	4.230	1.162	-7.734	-0.069
4		146.847	20.979	-1.989	-0.866	-14.133	-0.112
5		20.938	-106.143	-3.172	-2.294	5.252	0.113
6		63.187	15.759	0.622	4.974	2.075	-0.000
EfficientViT-Precision							
1	62.695	-128.788	-8.028	0.859	2.765	-1.316	0.008
2		-16.544	62.923	0.253	3.621	-12.244	0.166
3		-124.869	4.712	4.230	1.162	-8.155	-0.069
4		146.910	21.145	-1.989	-0.866	-13.233	-0.112
5		20.899	-106.007	-3.172	-2.294	4.992	0.112
6		63.198	15.761	0.622	4.974	2.005	-0.000
EfficientViT-Recall							
1	62.667	-128.816	-7.989	0.859	2.765	-0.928	0.008
2		-16.551	62.961	0.253	3.622	-12.896	0.166
3		-124.907	4.735	4.230	1.162	-8.707	-0.069
4		146.867	21.114	-1.989	-0.866	-11.932	-0.113
5		20.886	-106.115	-3.172	-2.294	4.889	0.112
6		63.170	15.741	0.622	4.974	1.814	-0.000
Model in Fig. 3 under EfficientViT-DJSCC-NoEqualization conditions							
EfficientViT-Accuracy							
1	-44.020	-38.843	275.129	1.924	1.978	47.147	0.000
2		-51.903	-84.152	-0.881	-1.047	93.905	0.000
3		66.244	-53.214	1.738	-0.279	71.377	0.000
4		-78.632	-35.214	0.458	2.515	34.492	0.000
5		-6.404	27.188	-1.103	-0.716	-337.171	0.000
6		64.367	-71.659	-1.634	-4.087	-108.441	0.000
EfficientViT-Precision							
1	-53.175	-56.343	286.785	1.926	1.980	45.960	0.000
2		-51.523	-70.393	-0.930	-1.090	105.882	0.000
3		62.920	-54.112	1.768	-0.255	72.418	0.000
4		-111.394	8.455	0.455	2.514	2.130	0.000
5		-41.345	42.701	-1.129	-0.738	-532.766	0.000
6		98.111	-85.710	-1.631	-4.084	-195.265	0.000
EfficientViT-Recall							
1	-46.460	-44.525	288.621	1.929	1.971	49.742	0.000
2		-50.752	-88.416	-0.867	-1.230	98.356	0.000
3		58.017	-58.214	1.726	-0.121	74.102	0.000
4		-89.070	-24.521	0.462	2.433	29.453	0.000
5		-12.033	58.197	-1.099	-1.622	-338.964	0.000
6		76.090	-73.528	-1.634	-4.001	-117.803	0.000
Model in Fig. 3 under EfficientViT-DJSCC-Equalization conditions							
EfficientViT-Accuracy							
1	78.623	12.773	79.416	1.843	1.401	73.357	0.000
2		-85.313	-77.315	-0.891	-0.934	57.183	0.000
3		-8.860	-30.991	1.733	-0.425	79.735	0.000
4		39.538	-45.119	0.456	2.447	22.400	0.000
5		3.132	21.592	-1.083	-0.656	-100.408	0.000
6		53.425	-0.220	-1.634	-4.083	23.030	0.000
EfficientViT-Precision							
1	71.093	21.850	71.994	1.842	1.380	78.662	0.000
2		-92.646	-76.294	-0.897	-1.106	62.805	0.000
3		-8.964	-29.287	1.729	-0.375	84.485	0.000
4		37.860	-24.102	0.464	2.488	35.706	0.000
5		3.253	18.605	-1.077	-0.584	-110.339	0.000
6		45.714	-2.757	-1.634	-4.083	22.709	0.000
EfficientViT-Recall							
1	76.957	13.288	82.881	1.844	1.401	68.630	0.000
2		-86.966	-78.347	-0.887	-0.933	54.429	0.000
3		-6.701	-31.646	1.736	-0.424	75.210	0.000
4		37.234	-42.442	0.460	2.449	29.623	0.000
5		2.833	23.648	-1.084	-0.656	-95.422	0.000
6		51.224	-1.029	-1.634	-4.083	18.436	0.000

TABLE XIII: PARAMETERS OF THE PROPOSED MODEL IN FIG. 4.

$i \backslash$ Parameters	μ_0	$\mu_{i,1}$	$\mu_{i,2}$	$\mu_{i,3}$	$\mu_{i,4}$	$\mu_{i,5} \times 10^3$	$\mu_{i,6}$
Model in Fig. 4 under MobileViT-DJSCC-AWGN conditions							
MobileViT-Accuracy							
1	62.684	-128.796	-7.953	0.859	2.765	-1.479	0.008
2		-16.485	63.021	0.253	3.622	-11.789	0.166
3		-124.913	4.799	4.230	1.162	-7.735	-0.069
4		146.848	20.979	-1.989	-0.866	-14.133	-0.112
5		20.940	-106.143	-3.172	-2.294	5.253	0.113
6		63.188	15.760	0.622	4.974	2.077	-0.000
MobileViT-Precision							
1	62.710	-128.773	-8.006	0.859	2.765	-0.859	0.008
2		-16.510	62.954	0.253	3.621	-13.243	0.166
3		-124.863	4.766	4.230	1.162	-8.821	-0.069
4		146.912	21.082	-1.989	-0.866	-11.727	-0.112
5		20.928	-106.037	-3.172	-2.294	4.095	0.112
6		63.213	15.777	0.622	4.974	1.795	-0.000
MobileViT-Recall							
1	62.704	-128.779	-7.983	0.859	2.765	-0.963	0.008
2		-16.508	62.973	0.253	3.621	-12.964	0.166
3		-124.873	4.785	4.230	1.162	-8.652	-0.069
4		146.900	21.058	-1.989	-0.866	-12.069	-0.112
5		20.927	-106.082	-3.172	-2.294	4.510	0.112
6		63.207	15.774	0.622	4.974	1.838	-0.000
Model in Fig. 4 under MobileViT-DJSCC-NoEqualization conditions							
MobileViT-Accuracy							
1	76.547	139.526	-2.260	-1.740	19.088	-346.541	0.000
2		91.702	-8.727	-0.996	1.584	76.653	0.000
3		160.336	4.079	-0.213	-7.216	-724.459	0.000
4		-243.199	12.691	-1.376	2.885	-518.624	0.000
5		-34.959	10.790	0.664	5.440	-335.882	0.000
6		-97.953	16.130	4.731	-0.619	82.356	0.000
MobileViT-Precision							
1	90.477	140.735	-1.055	-1.740	19.088	-438.509	0.000
2		102.485	-7.980	-0.980	1.586	84.079	0.000
3		141.563	4.019	-0.213	-7.216	-632.867	0.000
4		-240.220	12.201	-1.382	2.884	-611.749	0.000
5		-48.035	-1.321	0.664	5.440	-285.344	0.000
6		-112.681	15.272	4.750	-0.617	88.637	0.000
MobileViT-Recall							
1	131.058	242.503	100.634	-1.740	19.088	-421.929	0.000
2		129.363	-8.146	-0.998	1.606	84.912	0.000
3		60.489	3.643	-0.213	-7.215	-573.596	0.000
4		-313.567	10.194	-1.380	2.802	-513.195	0.000
5		-82.766	-29.084	0.664	5.455	-249.416	0.000
6		-168.380	18.853	4.737	-0.386	81.388	0.000
Model in Fig. 4 under MobileViT-DJSCC-Equalization conditions							
MobileViT-Accuracy							
1	65.556	15.372	100.237	1.857	1.125	97.250	0.000
2		-104.231	-63.269	-0.896	-1.251	77.814	0.000
3		-7.038	-50.052	1.778	0.104	106.866	0.000
4		27.374	3.578	0.469	2.590	76.312	0.000
5		6.988	21.071	-1.108	-0.593	-58.505	0.000
6		39.362	-6.100	-1.634	-4.081	12.551	0.000
MobileViT-Precision							
1	70.806	19.033	98.550	1.841	1.365	91.431	0.000
2		-97.600	-83.008	-0.916	-1.219	74.674	0.000
3		-8.464	-43.517	1.765	-0.134	98.479	0.000
4		33.617	-23.215	0.457	2.427	42.971	0.000
5		2.951	20.402	-1.090	-0.724	-111.654	0.000
6		44.789	-4.359	-1.634	-4.083	24.190	0.000
MobileViT-Recall							
1	61.338	-2.664	131.104	1.858	1.168	100.112	0.000
2		-82.141	-84.804	-0.899	-1.486	80.862	0.000
3		1.965	-73.355	1.779	0.273	107.026	0.000
4		23.807	-4.260	0.469	2.661	51.265	0.000
5		7.207	19.856	-1.108	-0.586	-73.983	0.000
6		34.188	-7.058	-1.634	-4.073	12.760	0.000

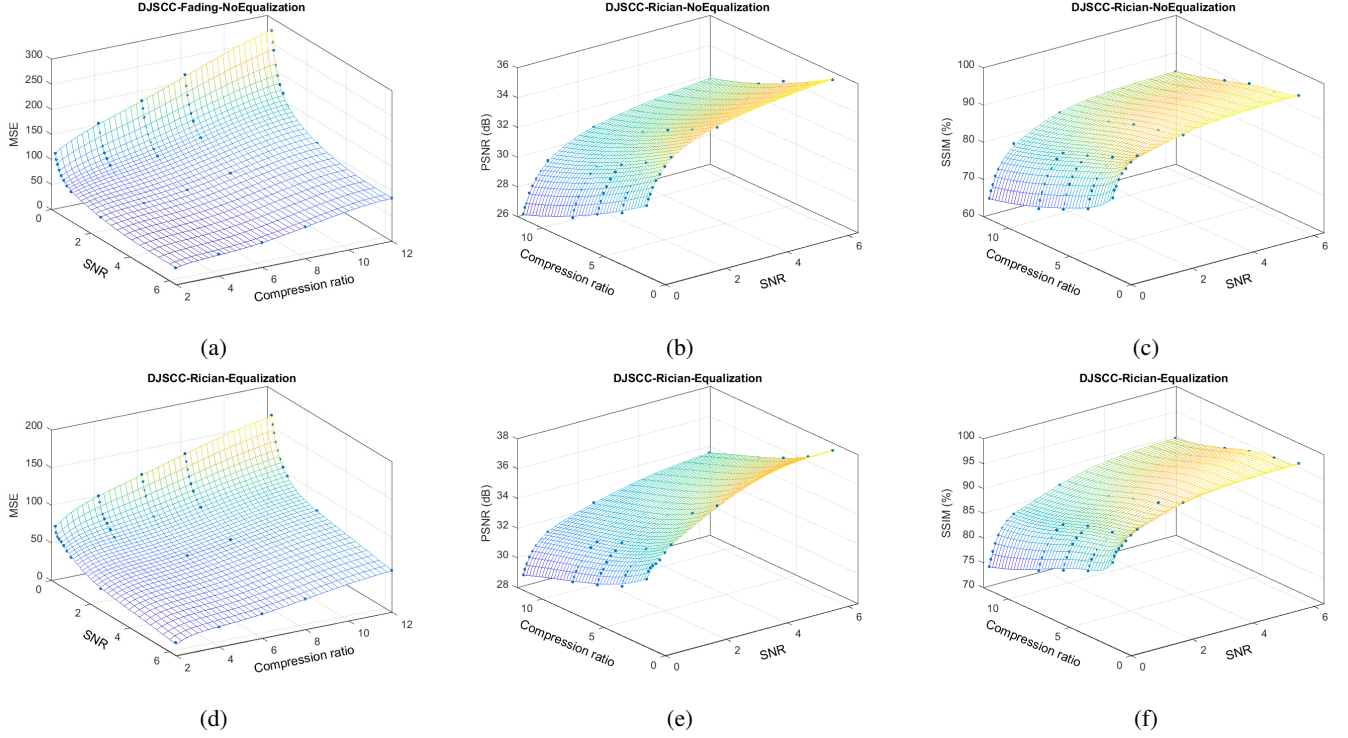


Fig. 8: Curve fitting models for the DJSCC case with Rician fading.

(a) No Equalization					
$\gamma \backslash \rho$	2	4	6	8	12
PSNR					
8 dB	35.5323	34.5115	33.8953	32.9589	31.7375
4 dB	34.4379	33.4852	32.7224	31.7944	30.6264
0 dB	33.2542	32.1079	31.2041	30.3017	29.2222
-1 dB	32.7513	31.5751	30.8529	29.9621	28.6026
-2 dB	32.3143	31.1285	30.4540	29.4272	28.2088
-3 dB	31.9693	30.8140	30.0337	29.0446	27.7857
-4 dB	31.5591	30.4281	29.6391	28.6420	27.3708
-5 dB	31.1633	30.1287	29.1439	28.2158	26.7033
-6 dB	30.8545	29.7005	28.7077	27.8178	26.4830
-7 dB	30.4555	29.1953	28.3460	27.3389	26.0702
SSIM					
8 dB	93.8425	91.7877	90.9587	88.7242	84.9465
4 dB	91.6769	89.7917	88.2895	85.9539	82.3632
0 dB	89.4480	85.9744	83.2750	80.9444	77.4457
-1 dB	88.1463	84.0082	82.0369	79.8207	74.2633
-2 dB	86.7859	82.4064	80.5796	77.6424	73.0812
-3 dB	85.7737	81.3526	78.7863	75.9628	71.6536
-4 dB	84.3830	79.6907	77.0092	74.0313	70.0491
-5 dB	82.7832	78.4813	74.3424	71.9268	67.9400
-6 dB	81.7795	76.3904	72.1920	70.0107	66.3748
-7 dB	79.8997	73.8710	70.6287	67.7948	64.4982
MSE					
8 dB	33.2201	43.0790	48.8054	62.7485	85.5222
4 dB	44.2726	54.4115	63.9038	79.9182	105.2756
0 dB	59.2468	78.9119	95.5641	114.5967	143.9191
-1 dB	65.8993	90.0280	104.2262	123.8198	164.1927
-2 dB	72.9877	99.4739	114.4857	139.0949	177.1709
-3 dB	79.4435	107.6998	126.7638	152.0240	192.7561
-4 dB	87.7928	118.8492	139.7152	167.2645	210.8477
-5 dB	97.5865	128.6889	156.9777	184.3022	238.5607
-6 dB	105.8723	142.3498	173.2149	201.3222	253.4160
-7 dB	117.0806	159.2828	187.2806	221.7056	275.1232

(b) Equalization					
$\gamma \backslash \rho$	2	4	6	8	12
PSNR					
8 dB	37.5434	36.3645	35.5251	34.7338	33.5640
4 dB	35.9322	34.6565	33.7365	33.1176	32.3238
0 dB	34.1393	33.0827	32.7458	32.0494	31.1866
-1 dB	33.7348	32.5586	32.3317	31.7800	30.7862
-2 dB	33.3296	32.2852	32.0889	31.3074	30.4894
-3 dB	33.1373	32.1823	31.7286	30.9917	30.1690
-4 dB	33.0693	31.9130	31.2998	30.6693	29.8311
-5 dB	32.9768	31.7191	31.0782	30.4606	29.4359
-6 dB	32.7755	31.3318	30.6733	30.1059	29.1004
-7 dB	32.2810	31.0378	30.3648	29.8058	28.7553
SSIM					
8 dB	96.0107	94.8403	93.8919	92.2208	89.6336
4 dB	94.4370	92.1631	90.5671	88.7956	86.5090
0 dB	91.5267	88.6594	87.6226	85.6246	83.1747
-1 dB	90.7046	87.5441	86.5827	84.8677	82.0500
-2 dB	89.8267	86.6873	85.8433	83.5439	80.9823
-3 dB	89.1070	85.9327	84.8314	82.3906	79.7572
-4 dB	88.6527	85.1783	83.4473	81.1974	78.2855
-5 dB	88.1347	84.3379	82.4638	80.1208	76.7694
-6 dB	87.3896	83.1672	81.1219	78.8675	75.3168
-7 dB	86.1955	82.1914	80.0279	77.6939	73.9138
MSE					
8 dB	17.2139	26.3414	32.1040	40.4677	55.2202
4 dB	28.9261	40.9717	49.8425	59.4779	74.8252
0 dB	46.7004	61.3167	66.5025	78.4057	96.6728
-1 dB	51.7015	68.8176	74.2685	84.7193	105.5506
-2 dB	56.4584	74.5385	79.1759	94.1022	113.8166
-3 dB	59.5376	78.4417	86.7114	102.0074	122.8543
-4 dB	61.2763	83.3804	95.7975	110.1298	132.8703
-5 dB	63.4603	88.2448	102.7553	118.3580	143.7088
-6 dB	67.7808	96.9349	112.7283	127.8829	153.9259
-7 dB	75.2589	104.2519	121.2174	137.3458	165.3323

TABLE XIV: DJSCC-Rician Empirical Results (Rician K-factor = 100).

TABLE XV: PARAMETERS OF THE PROPOSED MODEL IN FIG. 8.

$i \backslash$ Parameters	μ_0	$\mu_{i,1}$	$\mu_{i,2}$	$\mu_{i,3}$	$\mu_{i,4}$	$\mu_{i,5} \times 10^3$	$\mu_{i,6}$
Model in Fig. 8 under DJSCC-Rician-NoEqualization conditions							
Image-MSE							
1	125.422	-565.292	54.936	-1.496	-1.704	-462.412	0.000
2		1062.867	-19.406	1.180	-2.729	65.908	0.000
3		-237.212	-257.507	3.315	1.424	98.213	0.000
4		68.726	697.071	1.592	1.303	-301.644	0.000
5		962.544	-1949.674	1.942	1.888	-13.006	0.000
6		-66.970	152.779	3.226	-0.926	5.995	0.000
Image-PSNR							
1	11.913	52.632	-24.112	1.451	1.860	8.281	0.000
2		182.388	-0.352	-0.126	1.469	17.720	0.000
3		-227.264	6.142	1.429	-0.259	10.601	0.000
4		83.476	2.409	-0.411	0.413	-7.782	0.000
5		30.387	-18.387	-0.247	-1.171	1.363	0.000
6		-131.715	53.918	2.982	2.416	13.802	0.000
Image-SSIM							
1	25.252	132.102	-9.674	-1.740	19.149	-272.441	0.000
2		29.812	-16.978	-0.997	0.646	12.802	0.000
3		106.300	4.815	-0.213	-7.297	-621.515	0.000
4		-195.289	338.765	-1.382	-2.689	-407.316	0.000
5		-23.553	15.639	0.664	6.350	-348.237	0.000
6		-5.713	33.588	4.731	-0.242	-19.886	0.000
Model in Fig. 8 under DJSCC-Rician-Equalization conditions							
Image-MSE							
1	1381.339	-784.921	-1174.024	-2.900	-2.431	-603.699	0.000
2		-1015.634	981.981	2.762	0.233	-56.882	0.000
3		-541.429	81.469	-1.188	-0.641	-465.644	0.000
4		-549.150	225.128	-0.082	1.667	35.870	0.000
5		571.694	755.831	2.836	5.137	-474.889	0.000
6		1547.047	-2605.796	2.358	1.418	-31.094	0.000
Image-PSNR							
1	27.870	40.532	-5.429	0.633	-3.454	-126.943	0.000
2		-90.235	-4.694	0.156	-13.865	27.035	0.000
3		-24.079	97.139	2.658	3.058	28.234	0.000
4		-71.992	-13.919	-0.440	0.932	-72.389	0.000
5		-25.926	28.938	-0.269	20.030	-3.055	0.000
6		62.495	44.067	-2.346	-2.653	-22.026	0.000
Image-SSIM							
1	28.054	126.454	-15.319	-1.740	18.849	-316.707	0.000
2		57.594	-11.183	-0.996	1.969	-20.976	0.000
3		154.946	4.211	-0.213	-7.462	-769.663	0.000
4		-202.220	10.271	-1.375	3.975	-512.666	0.000
5		-36.332	7.585	0.663	3.825	-335.877	0.000
6		-2.733	9.592	4.730	-0.514	67.068	0.000