

# List of Accepted Papers

Following is the list of accepted ICASSP 2024 papers, sorted by paper title. You can use the search feature of your web browser to find your paper number. Notifications to all authors have also been sent by email. If you have not received your notification of the results by email, please contact us at [info@2024.ieeeicassp.org](mailto:info@2024.ieeeicassp.org).

Paper Number	Paper Title
1315	"IT IS OKAY TO BE UNCOMMON": QUANTIZING SOUND EVENT DETECTION NETWORKS ON HARDWARE ACCELERATORS WITH UNCOMMON SUB-BYTE SUPPORT
3377	1-D SPATIAL ATTENTION IN BINARIZED CONVOLUTIONAL NEURAL NETWORKS
4343	2D Human Pose Estimation Calibration and Keypoint Visibility Classification
9562	3D AUTOMATED QUANTITATIVE CALCULATIONS BASED ON CT IMAGES OF THE HIP JOINT
11925	3D CBCT CHALLENGE 2024: IMPROVED CONE BEAM CT RECONSTRUCTION USING SWINIR-BASED SINOGRAM AND IMAGE ENHANCEMENT
9987	3D Hand Joint and Grasping Estimation for Teleoperation System
1248	3-D Near-field Localization by Jointly Exploiting Spatial and Temporal Information Based on a Nonuniform Cross Array
5765	3D PARALLELISM FOR TRANSFORMERS VIA INTEGER PROGRAMMING
11565	3D PERCEPTUAL SOUNDFIELD RECONSTRUCTION VIA VIRTUAL MICROPHONE SYNTHESIS
10067	3D POINT CLOUD SEMANTIC SEGMENTATION BASED ON DIFFUSION MODEL
6588	3D POSE ESTIMATION FROM MONOCULAR VIDEO WITH CAMERA-BONE ANGLE REGULARIZATION ON THE IMAGE FEATURE
2858	3DSAM: SEGMENT ANYTHING IN NERF

<b>Paper Number</b>	<b>Paper Title</b>
4353	3M-TRANSFORMER: A MULTI-STAGE MULTI-STREAM MULTIMODAL TRANSFORMER FOR EMBODIED TURN-TAKING PREDICTION
2738	3S-TSE: EFFICIENT THREE-STAGE TARGET SPEAKER EXTRACTION FOR REAL-TIME AND LOW-RESOURCE APPLICATIONS
9281	6DOF SELD: SOUND EVENT LOCALIZATION AND DETECTION USING MICROPHONES AND MOTION TRACKING SENSORS ON SELF-MOTIONING HUMAN
10265	A 3D VIRTUAL TRY-ON METHOD WITH GLOBAL-LOCAL ALIGNMENT AND DIFFUSION MODEL
7550	A BAYESIAN APPROACH TO HIGH-ORDER LINK PREDICTION
4797	A BINARY BP DECODING USING POSTERIOR ADJUSTMENT FOR QUANTUM LDPC CODES
3674	A BI-PYRAMID MULTIMODAL FUSION METHOD FOR THE DIAGNOSIS OF BIPOLAR DISORDERS
8415	A BiRGAT Model for Multi-intent Spoken Language Understanding with Hierarchical Semantic Frames
11526	A BP Method for Track-Before-Detect
2465	A CCM-BASED JOINT DOA-FREQUENCY ESTIMATION AND SIGNAL RECOVERY WITH EFFICIENT SUB-NYQUIST SAMPLING
4570	A Chat About Boring Problems: Studying GPT-based text normalization
4413	A Closer Look at Wav2Vec2 Embeddings for On-device Single-channel Speech Enhancement
7266	A codec-based approach for video life-cycle characterization in social networks
8328	A COMPARATIVE ANALYSIS OF POETRY READING AUDIO: SINGING, NARRATING, OR SOMEWHERE IN BETWEEN?
6150	A COMPARATIVE STUDY ON ANNOTATION QUALITY OF CROWDSOURCING AND LLM VIA LABEL AGGREGATION
7551	A COMPARISON OF PARAMETER-EFFICIENT ASR DOMAIN ADAPTATION METHODS FOR UNIVERSAL SPEECH AND LANGUAGE MODELS

<b>Paper Number</b>	<b>Paper Title</b>
10309	A complete method for the 3D reconstruction of axonal pathways from 2 orthogonal 3D OCT images of the lamina cribrosa
1527	A COMPREHENSIVE ANALYSIS OF BIASES AND CUES IN NLU DATASETS AND MODELS WITH ICQ
4480	A COMPREHENSIVE FRAMEWORK FOR OCCLUDED HUMAN POSE ESTIMATION
4430	A COMPUTATIONALLY EFFICIENT SEMI-BLIND SOURCE SEPARATION APPROACH FOR NONLINEAR ECHO CANCELLATION BASED ON AN ELEMENT-WISE ITERATIVE SOURCE STEERING
8850	A CONCEPT FOR A SLAM BACK END HARDWARE ACCELERATOR
2992	A CONTRARIO PARADIGM FOR YOLO-BASED INFRARED SMALL TARGET DETECTION
4762	A CONVERGENT PRIMAL-DUAL DEEP PLUG-AND-PLAY ALGORITHM FOR CONSTRAINED IMAGE RESTORATION
9599	A Counterfactual Inspired Framework for Quantifying Edge Effects on GNNs Fairness
4905	A CROSS SEARCH METHOD FOR DATA AUGMENTATION IN NEURAL MACHINE TRANSLATION
1858	A crowdsourcing approach to video quality assessment
11463	A CTC ALIGNMENT-BASED NON-AUTOREGRESSIVE TRANSFORMER FOR END-TO-END AUTOMATIC SPEECH RECOGNITION
4646	A Deep Representation Learning-based Speech Enhancement Method Using Complex Convolution Recurrent Variational Autoencoder
6415	A DENSENET-BASED METHOD FOR DECODING AUDITORY SPATIAL ATTENTION WITH EEG
1378	A DENSITY-GUIDED TEMPORAL ATTENTION TRANSFORMER FOR INDISCERNIBLE OBJECT COUNTING IN UNDERWATER VIDEOS
7461	A DETAILED AUDIO-TEXT DATA SIMULATION PIPELINE USING SINGLE-EVENT SOUNDS

<b>Paper Number</b>	<b>Paper Title</b>
2700	A DISTRIBUTED JOINT INTEGRATED PROBABILISTIC DATA ASSOCIATION (JIPDA) FILTER WITH SOFT OBJECT ASSOCIATION
8136	A DUAL-PATH FRAMEWORK WITH FREQUENCY-AND-TIME EXCITED NETWORK FOR ANOMALOUS SOUND DETECTION
3263	A FACIAL EXPRESSION TRANSFER METHOD BASED ON 3DMM AND DIFFUSION MODELS
4860	A FAST BLIND DEBLURRING ALGORITHM USING LOCAL GRADIENT PRODUCT PRIOR
7284	A FAST, PERFORMANT, SECURE DISTRIBUTED TRAINING FRAMEWORK FOR LLM
7167	A FEDERATED GRAPH TO EMBEDDING APPROACH FOR KNOWLEDGE GRAPH COMPLETION
6498	A Fine-Grained Attribute Pre-labeling Method based on Label Dependency and Feature Similarity Dynamics
3175	A FINE-GRAINED TRI-MODAL INTERACTION MODEL FOR MULTIMODAL SENTIMENT ANALYSIS
5517	A FLEXIBLE ONLINE FRAMEWORK FOR PROJECTION-BASED STFT PHASE RETRIEVAL
11461	A FORMAT COMPLIANT ENCRYPTION METHOD FOR 3D OBJECTS ALLOWING HIERARCHICAL DECRYPTION
7959	A FOUNDATION MODEL FOR MUSIC INFORMATICS
5969	A FRAMEWORK FOR PORTRAIT STYLIZATION WITH SKIN-TONE AWARENESS AND NUDITY IDENTIFICATION
11929	A FULLBAND NEURAL NETWORK FOR AUDIO PACKET LOSS CONCEALMENT
6563	A fully differentiable model for unsupervised singing voice separation
5990	A GENERAL FRAMEWORK FOR ROTATION INVARIANT POINT CLOUD ANALYSIS
7292	A Generative Adversarial Framework for Dialogue Generation with Neural Architecture Search
7381	A GIBBS SAMPLER FOR BAYESIAN NONPARAMETRIC STATE-SPACE MODELS

<b>Paper Number</b>	<b>Paper Title</b>
9005	A GRAPH NEURAL NETWORK BASED APPROACH FOR FAULT DELINEATION IN SEISMIC DATA USING GRAPH TOTAL VARIATION AND MULTIGRAPH
1578	A GRAPH NEURAL NETWORK BASED FUSION OF MRI-DERIVED BRAIN NETWORK AND CLINICAL DATA FOR GLIOBLASTOMA SURVIVAL PREDICTION
8207	A GRAPH-PREDICTION-BASED APPROACH FOR DEBIASING UNDERREPORTED DATA
3273	A GREEN LEARNING APPROACH TO SPOOFED SPEECH DETECTION
3728	A GUIDED UPSAMPLING NETWORK FOR SHORT WAVE INFRARED IMAGES USING GRAPH REGULARIZATION
1469	A Hierarchical multi-proxy Loss with Dynamic Main-proxy for Deep Metric Learning
3912	A HYBRID CNN-TRANSFORMER FOR FOCAL LIVER LESION CLASSIFICATION
8486	A HYBRID DEEP-ONLINE LEARNING BASED METHOD FOR ACTIVE NOISE CONTROL IN WAVE DOMAIN
8251	A Hybrid Slow-time Coding Framework for Automotive MIMO Radar
2302	A JOINT DATA COMPRESSION AND TIME-DELAY ESTIMATION METHOD FOR DISTRIBUTED SYSTEMS VIA EXTREMUM ENCODING
4796	A JOINT LOOK ON LUNAR SATELLITE AND COOPERATIVE SURFACE PNT
1640	A KEYLESS EXTRACTION FRAMEWORK TARGETING AT DEEP LEARNING BASED IMAGE-WITHIN-IMAGE MODELS
3715	A Learning Resource Recommendation Algorithm Based on Online Learning Behavior
6868	A LEARNING-BASED MULTI-NODE FUSION POSITIONING METHOD USING WEARABLE INERTIAL SENSORS
3180	A LEARNING-BASED SYSTEM FOR AUTOMATIC INTENTIONAL NON-ADHERENCE DETECTION FROM DOSING VIDEOS
3481	A Lightweight Change Detection Method Based on Feature Interaction and Transformer for High Resolution Remote

<b>Paper Number</b>	<b>Paper Title</b>
	Sensing Images
10170	A LIGHTWEIGHT HYBRID MULTI-CHANNEL SPEECH EXTRACTION SYSTEM WITH DIRECTIONAL VOICE ACTIVITY DETECTION
2931	A LIGHT-WEIGHT STATE DETECTION MODEL FOR KALMAN-FILTER-BASED ACOUSTIC FEEDBACK CANCELLATION WITH RAPID RECOVERY FROM ABRUPT PATH CHANGES
8389	A LOW-LATENCY FFT-IFFT CASCADE ARCHITECTURE
8111	A Machine-Learning Model for Detecting Depression, Anxiety, and Stress from Speech
6005	A META-PRECONDITIONING APPROACH FOR DEEP Q-LEARNING
3105	A METHOD FOR BILEVEL OPTIMIZATION WITH CONVEX LOWER-LEVEL PROBLEM
3053	A METHOD FOR X-RAY IMAGE LANDMARKS LOCALIZATION USING CYCLIC COORDINATE-GUIDED STRATEGY
7492	A MODIFIED CRAMÉR-RAO BOUND FOR DISCRETE-TIME MARKOVIAN DYNAMIC SYSTEMS
11903	A MULTI-FILTER AND MULTI-SCALE U-NET FOR CONE-BEAM COMPUTED TOMOGRAPHY WITH HARDWARE CONSTRAINTS
3062	A MULTIMODAL APPROACH TO DEVICE-DIRECTED SPEECH DETECTION WITH LARGE LANGUAGE MODELS
5867	A MULTI-SCALE BIMODAL FUSION NETWORK FOR ROBUST AND ACCURATE ONLINE HANDWRITING RECOGNITION
4185	A MULTISCALE OBJECTIVE FUNCTION FOR CAMERA COLOR CORRECTION
4798	A NEAR-FIELD SOURCE LOCALIZATION METHOD FOR UNIFORM/SPARSE CENTRALLY SYMMETRIC RECTANGULAR ARRAYS
7410	A NEURAL SYNTAX PARSER FOR CORONARY ARTERY ANATOMICAL LABELING IN CORONARY CT ANGIOGRAPHY
7347	A Neurophysiological-Auditory "Listen Receipt" for Communication Enhancement

<b>Paper Number</b>	<b>Paper Title</b>
1565	A New Fourth-Order Sparse Array Generator Based on Sum-Difference Co-array Analysis
4557	A New Perspective on Understanding Resolution Limit via An Asymptotic Study of Christoffel-Darboux Kernel based Spectrum Estimator
9353	A New Pre-training Paradigm for Offline Multi-agent Reinforcement Learning with Suboptimal Data
5016	A new similarity-based relational knowledge distillation method
5623	A NOVEL 3-D FOCUSING SCHEME FOR DISTRIBUTED SAR TOMOGRAPHY
11910	A Novel Approach to WaveNet Architecture for RF Signal Separation with Learnable Dilation and Data Augmentation
7147	A NOVEL ARCHITECTURE OF DEEP FEATURE-BASED GAUSSIAN PROCESSES WITH AN ENSEMBLE OF KERNELS
7183	A NOVEL CASCADE INSTRUCTION TUNING METHOD FOR BIOMEDICAL NER
3358	A Novel Contrastive Diffusion Graph Convolutional Network for Few-Shot Skeleton-Based Action Recognition
10198	A NOVEL CROSS-SENSOR SELF-SUPERVISED LEARNING METHOD FOR ROTATING MACHINERY FAULT DIAGNOSIS
1907	A NOVEL DEMODULATION AND SELECTION PILOT POWER TRADE-OFF FOR CODEBOOK-BASED IRS WITH IMPERFECT CHANNEL ESTIMATES
8073	A NOVEL DISCRETE FRACTIONAL COMPLEX HADAMARD TRANSFORM FOR MEDICAL IMAGE ENCRYPTION
7081	A NOVEL ITERATIVE THRESHOLDING ALGORITHM FOR ARCTANGENT REGULARIZATION PROBLEM
5357	A NOVEL LOCAL-GLOBAL FEATURE FUSION FRAMEWORK FOR BODY-WEIGHT EXERCISE RECOGNITION WITH PRESSURE MAPPING SENSORS
6674	A NOVEL MEDICAL IMAGE FUSION FRAMEWORK INTEGRATING MULTI-SCALE ENCODER-DECODER WITH DISCRETE WAVELET DECOMPOSITION
9506	A Novel Multi-atlas Fusion Model Based On Contrastive Learning For Functional Connectivity Graph Diagnosis

<b>Paper Number</b>	<b>Paper Title</b>
8931	A NOVEL MULTIMODAL SENTIMENT ANALYSIS MODEL BASED ON GATED FUSION AND MULTI-TASK LEARNING
3150	A NOVEL RESIDUAL-GUIDED LEARNING METHOD FOR IMAGE STEGANOGRAPHY
5134	A One-Class Approach to Detect Super-Resolution Satellite Imagery with Spectral Features
8638	A parameterized generative adversarial network using cyclic projection for explainable medical image classifications
6132	A PLS-INTEGRATED LASSO METHOD WITH APPLICATION IN INDEX TRACKING
3216	A PRACTICAL ONLINE MULTICHANNEL DEREVERBERATION APPROACH WITH DATA-REUSE TECHNIQUE
7041	A PRIOR DRIVEN SEMI-SUPERVISED VITGAN FOR IMAGE RECOLORIZATION
2226	A PROBABILITY GRADIENT BASED APPROACH FOR SAMPLING BOUNDARIES OF IN-DOMAIN DATA
4564	A Prompt-based Method With Multi-View Optimization for Open Relation Extraction
10384	A Property-Guided Diffusion Model for Generating Molecular Graphs
5455	A RAY-TRACING BASED FINGERPRINTING METHOD FOR PASSIVE LOCALIZATION IN URBAN NLOS ENVIRONMENT
8948	A REAL-TIME ACTIVE SPEAKER DETECTION SYSTEM INTEGRATING AN AUDIO-VISUAL SIGNAL WITH A SPATIAL QUERYING MECHANISM
9012	A REAL-TIME LYRICS ALIGNMENT SYSTEM USING CHROMA AND PHONETIC FEATURES FOR CLASSICAL VOCAL PERFORMANCE
7076	A REAL-TIME VIDEO QUALITY METRIC FOR HTTP ADAPTIVE STREAMING
3574	A RECONSTRUCTION-BASED FEATURE ADAPTATION FOR ANOMALY DETECTION WITH SELF-SUPERVISED MULTI-SCALE AGGREGATION



<b>Paper Number</b>	<b>Paper Title</b>
2450	A Reduced-Reference Quality Assessment Metric for Textured Mesh Digital Humans
1463	A Relation-Aware Heterogeneous Graph Transformer on Dynamic Fusion for Multimodal Classification Tasks
8720	A RIEMANNIAN-BASED JOINT DESIGN FRAMEWORK OF MIMO RADAR TRANSMIT WAVEFORM AND RECEIVE FILTER VIA INFORMATION THEORY
4345	A ROBUST AND SCALABLE METHOD WITH AN ANALYTIC SOLUTION FOR MULTI-SUBJECT FMRI DATA ANALYSIS
5643	A ROBUST AUDIO DEEPPAKE DETECTION SYSTEM VIA MULTI-VIEW FEATURE
11503	A ROBUST FRAMEWORK TO DESIGN OPTIMAL SENSOR LOCATIONS FOR TOA OR RSS SOURCE LOCALIZATION TECHNIQUES
4076	A Robust GLRT Detector against Missing Data in Cooperative Sensing
8800	A ROBUST PITCH-FUSION MODEL FOR SPEECH EMOTION RECOGNITION IN TONAL LANGUAGES
4433	A ROBUST QUANTILE HUBER LOSS WITH INTERPRETABLE PARAMETER ADJUSTMENT IN DISTRIBUTIONAL REINFORCEMENT LEARNING
2018	A SALIENCY ENHANCED FEATURE FUSION BASED MULTISCALE RGB-D SALIENT OBJECT DETECTION NETWORK
7436	A SCALABLE SPARSE TRANSFORMER MODEL FOR SINGING MELODY EXTRACTION
11921	A SELF-SUPERVISED LEARNING APPROACH FOR DETECTING NON-PSYCHOTIC RELAPSES USING WEARABLE-BASED DIGITAL PHENOTYPING
3439	A SELF-SUPERVISED PRESSURE MAP HUMAN KEYPOINT DETECTION APPROCH: OPTIMIZING GENERALIZATION AND COMPUTATIONAL EFFICIENCY ACROSS DATASETS
9120	A SEPARATION PRIORITY PIPELINE FOR SINGLE-CHANNEL SPEECH SEPARATION IN NOISY ENVIRONMENTS
2383	A SEQUENTIAL AVERAGING PLUG-AND-PLAY METHOD FOR IMAGE RESTORATION VIA FIXED-POINT PROJECTION

<b>Paper Number</b>	<b>Paper Title</b>
10168	A SIMPLE AND EFFECTIVE METHOD FOR ANOMALY DETECTION ON ATTRIBUTED GRAPHS VIA FEATURE CONSISTENCY
2801	A Smoothed Bregman Proximal Gradient Algorithm for Decentralized Nonconvex Optimization
1461	A SOFT CONTRASTIVE LEARNING-BASED PROMPT MODEL FOR FEW-SHOT SENTIMENT ANALYSIS
8613	A SOUND APPROACH: USING LARGE LANGUAGE MODELS TO GENERATE AUDIO DESCRIPTIONS FOR EGOCENTRIC TEXT-AUDIO RETRIEVAL
8645	A SPATIAL LONG-TERM ITERATIVE MASK ESTIMATION APPROACH FOR MULTI-CHANNEL SPEAKER DIARIZATION AND SPEECH RECOGNITION
2161	A SPEAKER RECOGNITION METHOD BASED ON STABLE LEARNING
10394	A SPECTRAL ANALYSIS OF GRAPH NEURAL NETWORKS ON DENSE AND SPARSE GRAPHS
2824	A STATISTICAL CHARACTERIZATION OF COMMUNICATION PERFORMANCE IN RIS-AIDED NETWORKS
7297	A STEERED RESPONSE POWER APPROACH WITH BILINEAR PREDICTION-BASED TRADE-OFF PREWHITENING FOR SPEAKER LOCALIZATION
1690	A STOCHASTIC GRADIENT APPROACH FOR COMMUNICATION EFFICIENT CONFEDERATED LEARNING
2770	A Stochastic Proximal WMMSE for Ergodic Sum Rate Maximization
10275	A STUDY OF MISPRONUNCIATION DETECTION AND DIAGNOSIS BASED ON META-LEARNING
2806	A STUDY OF MULTICHANNEL SPATIOTEMPORAL FEATURES AND KNOWLEDGE DISTILLATION ON ROBUST TARGET SPEAKER EXTRACTION
3187	A STUDY ON COMBINING NON-PARALLEL AND PARALLEL METHODOLOGIES FOR MANDARIN-ENGLISH CROSS-LINGUAL VOICE CONVERSION
4187	A STUDY ON GRAPH EMBEDDING FOR SPEAKER RECOGNITION

<b>Paper Number</b>	<b>Paper Title</b>
2325	A STUDY ON THE ADVERSE IMPACT OF SYNTHETIC SPEECH ON SPEECH RECOGNITION
9834	A Supervised Information Enhanced Multi-granularity Contrastive Learning Framework for EEG based Emotion Recognition
6974	A TARGETED ADVERSARIAL ATTACK METHOD FOR MULTI-CLASSIFICATION MALICIOUS TRAFFIC DETECTION
11909	A TIME-FREQUENCY BAND-SPLIT NEURAL NETWORK FOR REAL-TIME FULL-BAND PACKET LOSS CONCEALMENT
4295	A TRANSFORMER APPROACH FOR POLYPHONIC AUDIO-TO-SCORE TRANSCRIPTION
1566	A TRI-DYNAMIC PREPROCESSING FRAMEWORK FOR UGC VIDEO COMPRESSION
5109	A TWO-STAGE DEHAZING FRAMEWORK BASED ON INVERTED IMAGE CURVE-ENHANCEMENT
9417	A TWO-STAGE FRAMEWORK IN CROSS-SPECTRUM DOMAIN FOR REAL-TIME SPEECH ENHANCEMENT
11917	A U-NET ARCHITECTURE FOR TIME-FREQUENCY INTERFERENCE SIGNAL SEPARATION OF RF WAVEFORMS
10178	A UNIFIED DNN-BASED SYSTEM FOR INDUSTRIAL PIPELINE SEGMENTATION
1226	A UNIFIED FRAMEWORK FOR MULTI-INTENT SPOKEN LANGUAGE UNDERSTANDING WITH PROMPTING
2091	A UNIFIED FRONT-END FRAMEWORK FOR ENGLISH TEXT-TO-SPEECH SYNTHESIS
7036	A UNIFIED LOSS FUNCTION TO TACKLE INTER-CLASS AND INTRA-CLASS DATA IMBALANCE IN SOUND EVENT DETECTION
1047	A VARIABLE SMOOTHING FOR NONCONVEXLY CONSTRAINED NONSMOOTH OPTIMIZATION WITH APPLICATION TO SPARSE SPECTRAL CLUSTERING
7832	A WASSERSTEIN GRAPH DISTANCE BASED ON DISTRIBUTIONS OF PROBABILISTIC NODE EMBEDDINGS
8952	A weighted-variance variational autoencoder model for speech enhancement

<b>Paper Number</b>	<b>Paper Title</b>
8961	AAT: ADAPTING AUDIO TRANSFORMER FOR VARIOUS ACOUSTICS RECOGNITION TASKS
11506	Absolute Security in Terahertz Wireless Links
3666	ACCELERATED RECOVERY OF SPECTRALLY SPARSE SIGNALS VIA MODIFIED PROXIMAL GRADIENT IN HANKEL SPACE
9936	ACCELERATING GRADIENT DESCENT FOR OVER-PARAMETERIZED ASYMMETRIC LOW-RANK MATRIX SENSING VIA PRECONDITIONING
2107	ACCENT-SPECIFIC VECTOR QUANTIZATION FOR JOINT UNSUPERVISED AND SUPERVISED TRAINING IN ACCENT ROBUST SPEECH RECOGNITION
9373	Accurate and Robust Scene Text Recognition via Adversarial Training
6057	ACCURATE GIGAPIXEL CROWD COUNTING BY ITERATIVE ZOOMING AND REFINEMENT
8578	ACCURATE INTERPOLATION OF SCATTERED DATA VIA LEARNING RELATION GRAPH
7271	ACOUSTIC BPE FOR SPEECH GENERATION WITH DISCRETE TOKENS
9014	ACTIVATION COMPRESSION OF GRAPH NEURAL NETWORKS USING BLOCK-WISE QUANTIZATION WITH IMPROVED VARIANCE MINIMIZATION
2333	ACTIVE EXPLAINABLE RECOMMENDATION WITH LIMITED LABELING BUDGETS
5967	ACTIVE LEARNING FOR SOUND EVENT CLASSIFICATION USING BAYESIAN NEURAL NETWORKS WITH GAUSSIAN VARIATIONAL POSTERIOR
2933	ACTIVE LEARNING WITH CORE-SET SAMPLING AND SCALE-SENSITIVE LOSS FOR 3D OBJECT DETECTION
7978	ACTIVE NOISE CONTROL OVER 3D SPACE WITH A DYNAMIC NOISE SOURCE
4052	ACTIVE NOISE CONTROL OVER A LARGE REGION WITH MULTIPLE SPHERICAL MICROPHONE ARRAYS IN WAVE DOMAIN

<b>Paper Number</b>	<b>Paper Title</b>
9895	ACTIVITY RECOGNITION METHOD BASED ON KERNEL SUPERVISED LAPLACIAN EIGENMAPS
6794	ADAFI: ADAPTIVE CLIENT SELECTION AND DYNAMIC CONTRIBUTION EVALUATION FOR EFFICIENT FEDERATED LEARNING
9832	ADAMER-CTC: CONNECTIONIST TEMPORAL CLASSIFICATION WITH ADAPTIVE MAXIMUM ENTROPY REGULARIZATION FOR AUTOMATIC SPEECH RECOGNITION
1820	AdaPlus: Integrating Nesterov Momentum and Precise Stepsize Adjustment on AdamW Basis
5503	ADAPTER-BASED INCREMENTAL LEARNING FOR FACE FORGERY DETECTION
8853	ADAPTING FRECHET AUDIO DISTANCE FOR GENERATIVE MUSIC EVALUATION
4814	ADAPTING LARGE LANGUAGE MODEL WITH SPEECH FOR FULLY FORMATTED END-TO-END SPEECH RECOGNITION
6768	ADAPTING PITCH-BASED SELF SUPERVISED LEARNING MODELS FOR TEMPO ESTIMATION
4508	Adaptive Chroma Block Vector Derivation From Luma for Screen Content Coding
1939	Adaptive Confidence Multi-View Hashing for Multimedia Retrieval
5188	ADAPTIVE DATA AUGMENTATION FOR ASPECT SENTIMENT QUAD PREDICTION
3306	ADAPTIVE FOURIER DECOMPOSITION BASED SIGNAL EXTRACTION ON WEAK ELECTROMAGNETIC FIELD
10109	Adaptive Gaussian Regularization Constrained Sparse Subspace Clustering for Image Segmentation
1537	ADAPTIVE GRID 2-D DIRECTION OF ARRIVAL ESTIMATION METHOD USING AN INTEGRATED DICTIONARY
4698	ADAPTIVE HEAD POSE ESTIMATION WITH REAL-TIME STRUCTURED LIGHT
4588	ADAPTIVE IMAGE-ENHANCED KNOWLEDGE GRAPH COMPLETION
7628	ADAPTIVE JOINT CHANNEL ESTIMATION/DATA DETECTION IN FLEXIBLE MULTICARRIER MIMO SYSTEMS -

<b>Paper Number</b>	<b>Paper Title</b>
	A TENSOR-BASED APPROACH
3959	ADAPTIVE KALMANNET: DATA-DRIVEN KALMAN FILTER WITH FAST ADAPTATION
2075	ADAPTIVE MULTI-ARMED BANDIT LEARNING FOR TASK OFFLOADING IN MOBILE EDGE COMPUTING
8653	Adaptive Multi-Exposure Fusion for Enhanced Neural Radiance Fields
7636	ADAPTIVE MULTIVIEW COMMUNITY-PRESERVED GRAPH CONVOLUTIONAL NETWORK FOR MULTIATLAS-BASED FUNCTIONAL CONNECTIVITY ANALYSIS
3672	Adaptive Multi-View Joint Contrastive Learning on Graphs
10002	ADAPTIVE ORDER AGGREGATOR AND EXTRACTOR GRAPH NEURAL NETWORK
4135	Adaptive parameter sharing for multi-agent reinforcement learning
8588	ADAPTIVE PEDESTRIAN TRAJECTORY PREDICTION VIA TARGET-DIRECTED ANGLE AUGMENTATION
1621	ADAPTIVE PROMPT CONSTRUCTION METHOD FOR RELATION EXTRACTION
7010	ADAPTIVE QUANTIZATION WITH MIXED-PRECISION BASED ON LOW-COST PROXY
7844	Adaptive Reweighted Sparse Belief Propagation Decoding for Polar Codes
8655	ADAPTIVE SECONDARY TRANSFORM SETS FOR VIDEO CODING BEYOND AV1
5443	Adaptive Sensor Selection With Deterministic Priors for DoA Tracking
8681	ADAPTIVE SPATIAL-TEMPORAL HYPERGRAPH FUSION LEARNING FOR NEXT POI RECOMMENDATION
5017	ADAPTIVE SPEECH EMOTION REPRESENTATION LEARNING BASED ON DYNAMIC GRAPH
8672	Adaptive Super Resolution For One-Shot Talking-Head Generation
3635	ADAPTIVE VIDEO WATERMARKING WITH PERCEPTUAL GUARANTEE AND EFFICIENCY OPTIMIZATION

<b>Paper Number</b>	<b>Paper Title</b>
7216	ADAPTIVE-AVG-POOLING BASED ATTENTION VISION TRANSFORMER FOR FACE ANTI-SPOOFING
8313	ADDRESSING CONFOUNDS IN FUNCTIONAL CONNECTIVITY ANALYSES OF CALCIUM IMAGING
7480	Addressing Data Scarcity In Voice Disorder Detection with Self-Supervised Models
5478	ADHD DIAGNOSIS AND BIOMARKER DETECTION BASED ON MULTIMODAL GRAPH CONVOLUTIONAL NEURAL NETWORK
2191	ADIFT: ZERO-SHOT GENERATIVE MODEL ADAPTION VIA ADAPTIVE DOMAIN-INVARIANT FEATURE TRANSFER
4307	Advancing Acoustic Howling Suppression through Recursive Training of Neural Networks
11911	ADVANCING THE FRONTIERS OF DEEP LEARNING FOR LOW-DOSE 3D CONE-BEAM COMPUTED TOMOGRAPHY (CT) RECONSTRUCTION
11534	Adversarial Continual Learning to Transfer Self-Supervised Speech Representations for Voice Pathology Detection
8316	Adversarial Domain Adaptation for Classification with Nested Dichotomies
9652	ADVERSARIAL JAMMING FOR AUTOENCODER DISTRIBUTION MATCHING
4633	ADVERSARIAL LEARNING ON COMPRESSED POSTERIOR SPACE FOR NON-ITERATIVE SCORE-BASED END-TO-END TEXT-TO-SPEECH
9094	Adversarial Robustness of Convolutional Models Learned in the Frequency Domain
6368	ADVERSARIAL SPEECH FOR VOICE PRIVACY PROTECTION FROM PERSONALIZED SPEECH GENERATION
4062	ADVSHADOW: EVADING DEEPFAKE DETECTION VIA ADVERSARIAL SHADOW ATTACK
3141	ADVSV: AN OVER-THE-AIR ADVERSARIAL ATTACK DATASET FOR SPEAKER VERIFICATION
8559	AdvTTS: Adversarial Text-to-Speech Synthesis Attack on Speaker Identification Systems

<b>Paper Number</b>	<b>Paper Title</b>
8708	AEAM3D: ADVERSE ENVIRONMENT-ADAPTIVE MONOCULAR 3D OBJECT DETECTION VIA FEATURE EXTRACTION REGULARIZATION
8391	AEGIS-Net: Attention-guided Multi-Level Feature Aggregation for Indoor Place Recognition
8790	Aerial-IRS-Assisted Load Balancing in Downlink Networks
7713	AGADIR: Towards Array-Geometry Agnostic Directional Speech Recognition
2588	AHRNet: Attention and Heatmap-based Regressor for Hand Pose Estimation and Mesh Recovery
7621	AINUR: HARMONIZING SPEED AND QUALITY IN DEEP MUSIC GENERATION THROUGH LYRICS-AUDIO EMBEDDINGS
4678	ALIGN, ADAPT AND INJECT: AUDIO-GUIDED IMAGE GENERATION, EDITING AND STYLIZATION
3362	All Neural Kronecker Product Beamforming for Speech Extraction with Large-scale Microphone Arrays
5296	ALLEVIATING HALLUCINATIONS VIA SUPPORTIVE WINDOW INDEXING IN ABSTRACTIVE SUMMARIZATION
3709	AlphaRotate: A Rotation Detection Benchmark using TensorFlow
11535	ALTERNATING LEAST-SQUARES-BASED MICROPHONE ARRAY PARAMETER ESTIMATION FOR A SINGLE-SOURCE REVERBERANT AND NOISY ACOUSTIC SCENARIO
4142	AMBISONICS NETWORKS - THE EFFECT OF RADIAL FUNCTIONS REGULARIZATION
6831	AN ACCURATE AND EFFICIENT NEURAL NETWORK FOR OCTA VESSEL SEGMENTATION AND A NEW DATASET
3235	AN ACTIVE NOISE CONTROL SYSTEM BASED ON SOUNDFIELD INTERPOLATION USING A PHYSICS-INFORMED NEURAL NETWORK
3826	AN ADAPTER-BASED UNIFIED MODEL FOR MULTIPLE SPOKEN LANGUAGE PROCESSING TASKS
7430	AN ADAPTIVE ALGORITHM FOR TRACKING THIRD-ORDER COUPLED CANONICAL POLYADIC DECOMPOSITION



<b>Paper Number</b>	<b>Paper Title</b>
8738	AN ANCHOR LEARNING APPROACH FOR CITATION FIELD LEARNING
10024	An Asymptotically Achievable Rate Bound for Establishing High-Fidelity Entanglements in Quantum Networks
9668	AN ATTENTION-ENHANCED RETENTIVE BROAD LEARNING SYSTEM FOR SUBJECT-GENERIC EMOTION RECOGNITION FROM EEG SIGNALS
11874	AN AUDIO-QUALITY-BASED MULTI-STRATEGY APPROACH FOR TARGET SPEAKER EXTRACTION IN THE MISP 2023 CHALLENGE
8453	AN AUDIO-TEXTUAL DIFFUSION MODEL FOR CONVERTING SPEECH SIGNALS INTO ULTRASOUND TONGUE IMAGING DATA
5497	AN EFFECTIVE MIXTURE-OF-EXPERTS APPROACH FOR CODE-SWITCHING SPEECH RECOGNITION LEVERAGING ENCODER DISENTANGLEMENT
8216	AN EFFICIENT ALGORITHM FOR CLUSTERED MULTI-TASK COMPRESSIVE SENSING
5849	AN EFFICIENT ALGORITHM FOR MULTIUSER SUM-RATE MAXIMIZATION OF LARGE-SCALE Active RIS-AIDED MIMO SYSTEM
8339	AN EFFICIENT ALTERNATING RIEMANNIAN/PROJECTED GRADIENT DESCENT ASCENT ALGORITHM FOR FAIR PRINCIPAL COMPONENT ANALYSIS
3158	An Efficient and Interpretable Speech Enhancement Network via Deep Dictionary Learning
8763	AN EFFICIENT HIERARCHICAL BLOCK COORDINATE DESCENT METHOD FOR TIME-VARYING GRAPHICAL LASSO
6921	An Efficient Temporary Deepfake Location Approach Based Embeddings for Partially Spoofed Audio Detection
9411	An Efficient Transformer for Demosaicing via Compressed Multi-branch Attention Mechanism
1383	AN EMPIRICAL INVESTIGATION OF DOMAIN ADAPTATION ABILITY FOR CHINESE SPELLING CHECK MODELS
7060	AN EMPIRICAL STUDY ON THE IMPACT OF POSITIONAL ENCODING IN TRANSFORMER-BASED MONAURAL SPEECH

<b>Paper Number</b>	<b>Paper Title</b>
	ENHANCEMENT
1955	AN END-TO-END EEG CHANNEL SELECTION METHOD WITH RESIDUAL GUMBEL SOFTMAX FOR BRAIN-ASSISTED SPEECH ENHANCEMENT
3887	AN ERROR SELF-CORRECTED DOA ESTIMATION MODEL FOR SPARSE ARRAY BASED ON ANM
7610	An Experimental Comparison Of Multi-view Self-supervised Methods For Music Tagging
6584	AN EXPERIMENTAL COMPARISON OF NOISE-ROBUST TEXT-TO-SPEECH SYNTHESIS SYSTEMS BASED ON SELF-SUPERVISED REPRESENTATION
3484	AN EXPLAINABLE PROXY MODEL FOR MULTILABEL AUDIO SEGMENTATION
7170	AN EXPLICIT MULTI-MODAL FUSION METHOD FOR SIGN LANGUAGE TRANSLATION
4004	AN INITIAL INVESTIGATION OF NEURAL REPLAY SIMULATOR FOR OVER-THE-AIR ADVERSARIAL PERTURBATIONS TO AUTOMATIC SPEAKER VERIFICATION
7379	AN INTERPRETABLE AND GENERALIZABLE SPEECH DETECTOR BASED ON A CNN-LSTM FRAMEWORK
7025	AN INVESTIGATION OF DISTRIBUTION ALIGNMENT IN MULTI-GENRE SPEAKER RECOGNITION
3126	AN MVDR-EMBEDDED U-NET BEAMFORMER FOR EFFECTIVE AND ROBUST MULTICHANNEL SPEECH ENHANCEMENT
6904	AN OPTIMIZED INTERLEAVED OFDM CHIRP ORTHOGONAL WAVEFORM DESIGN FOR DECHIRPED MINIATURE MMW MIMO RADAR
9579	AN UNSUPERVISED SEGMENTATION OF VOCAL BREATH SOUNDS
10434	Analysis and Utilization of Hidden Information in Model Inversion Attacks
1473	Analysis of An Elliptic Localization Algorithm Using Fixed Point Iteration
7938	ANALYSIS OF HIGH-ORDER BRAIN NETWORKS RESOLVED IN TIME AND FREQUENCY USING CP DECOMPOSITION

<b>Paper Number</b>	<b>Paper Title</b>
7397	Analysis of the Memorization and Generalization Capabilities of AI Agents: Are Continual Learners Robust?
3991	ANALYSIS OF THE SINR IN LEO-PNT SYSTEMS WITH 5G PRS MULTIPLEXING: INTEGRATION OF PRS AND NTN
10440	Analytical performance assessment of 2-D Tensor ESPRIT in terms of physical parameters
10419	ANALYZING ADVERSARIAL VULNERABILITIES OF GRAPH LOTTERY TICKETS
6902	ANCHOR-GUIDED GAN WITH CONTRASTIVE LOSS FOR LOW-RESOURCE OUT-OF-DOMAIN DETECTION
7368	ANIM-400K: A LARGE-SCALE DATASET FOR AUTOMATED END TO END DUBBING OF VIDEO
5084	ANM-BASED SOURCE LOCALIZATION UNDER MIXED FIELD
9884	ANOMALOUS SOUND DETECTION BY FEATURE-LEVEL ANOMALY SIMULATION
2734	ANOMALY DETECTION FROM A FREQUENCY PERSPECTIVE: M-BAND WAVELET PACKET ANOMALY DETECTION NETWORK
1886	ANOMALY-AWARE SEMANTIC SELF-ALIGNMENT FRAMEWORK FOR VIDEO-BASED PERSON RE-IDENTIFICATION
9025	ANONYMIZING SPEAKER VOICES: EASY TO IMITATE, DIFFICULT TO RECOGNIZE?
3373	ANTI-DECEPTION JAMMING POWER OPTIMIZATION STRATEGY FOR MULTI-TARGET TRACKING TASKS IN MULTI-RADAR SYSTEMS
7517	APOLLO'S UNHEARD VOICES: GRAPH ATTENTION NETWORKS FOR SPEAKER DIARIZATION AND CLUSTERING FOR FEARLESS STEPS APOLLO COLLECTION
9190	APPLICATION OF SNNs MODEL BASED ON MULTI-DIMENSIONAL ATTENTION IN DRONE RADIO FREQUENCY SIGNAL CLASSIFICATION
7243	APPLYING HYBRID QUANTUM LSTM FOR INDOOR LOCALIZATION BASED ON RSSI
4434	AQF: Assessing the Quality of Hyperspectral Reconstruction with a Learnable Metric

<b>Paper Number</b>	<b>Paper Title</b>
3229	ARBITRARY STYLE TRANSFER BASED ON CONTENT INTEGRITY AND STYLE CONSISTENCY ENHANCEMENT
2796	Arbitrary Style Transfer with Prototype-based Channel Alignment
4006	ARCHITECTURE-AGNOSTIC ITERATIVE BLACK-BOX CERTIFIED DEFENSE AGAINST ADVERSARIAL PATCHES
9279	ARE DEEP NEURAL NETWORKS ROBUST TO NAMED ENTITIES? AN ADVERSARIAL ATTACK AND DEFENSE PERSPECTIVE
7809	ARE SNNs TRULY ENERGY-EFFICIENT? - A HARDWARE PERSPECTIVE
2735	ARE SOFT PROMPTS GOOD ZERO-SHOT LEARNERS FOR SPEECH RECOGNITION?
3477	ARFA: AN ASYMMETRIC RECEPTIVE FIELD AUTOENCODER MODEL FOR SPATIOTEMPORAL PREDICTION
3705	ARRAY GEOMETRY OPTIMIZATION FOR REGION-OF-INTEREST NEAR-FIELD BEAMFORMING
3839	ASFORMER: LEARNING FROM ADJACENT SCALE
2595	ASPED: AN AUDIO DATASET FOR DETECTING PEDESTRIANS
6798	AS-PVAD: A FRAME-WISE PERSONALIZED VOICE ACTIVITY DETECTION NETWORK WITH ATTENTIVE SCORE LOSS
4129	ASSESSING GNSS CARRIER-TO-NOISE-DENSITY RATIO ESTIMATION IN THE PRESENCE OF MEACONER INTERFERENCE
6255	ASSESSING VIBROACOUSTIC SOUND MESSAGE THROUGH THE BIOSIGNAL OF HUMAN SPEECH: EVIDENCE OF IMPROVED WELLBEING
4914	ASYMMETRIC CLEAN SEGMENTS-GUIDED SELF-SUPERVISED LEARNING FOR ROBUST SPEAKER VERIFICATION
7905	Asymptotic Behavior of Super-resolution Sparse Bayesian Learning
9820	ASYMPTOTICALLY TIGHT MISSPECIFIED BAYESIAN CRAMÉR-RAO BOUND

<b>Paper Number</b>	<b>Paper Title</b>
9715	ASYNCHRONOUS DIFFUSION LEARNING WITH AGENT SUBSAMPLING AND LOCAL UPDATES
4073	ATTA-NET: ATTENTION AGGREGATION NETWORK FOR AUDIO-VISUAL EMOTION RECOGNITION
1968	ATTENTION DECOUPLING FOR QUERY-BASED OBJECT DETECTION
8911	ATTENTION IS ALL YOU NEED FOR BLIND ROOM VOLUME ESTIMATION
5174	ATTENTION-BASED SPATIAL-FREQUENCY INFORMATION NETWORK FOR UNDERWATER SINGLE IMAGE SUPER-RESOLUTION
5409	ATTENTION-DRIVEN MULTICHANNEL SPEECH ENHANCEMENT IN MOVING SOUND SOURCE SCENARIOS
2297	ATTENTION-GUIDED ADAPTATION FOR CODE-SWITCHING SPEECH RECOGNITION
3689	ATTENTIONLUT: ATTENTION FUSION-BASED CANONICAL POLYADIC LUT FOR REAL-TIME IMAGE ENHANCEMENT
7925	AttHear: Explaining Audio Transformers Using Attention-Aware NMF
9405	ATTRIBUTE-AWARE AMPLIFICATION OF FACIAL FEATURE SEQUENCES FOR FACIAL EMOTION RECOGNITION
1767	Attribute-aware Head Swapping Guided by 3D Modeling
3185	Attribution-based Scanline Perturbation Attack on 3D Detectors of LiDAR Point Clouds
5085	ATTR-INT: A SIMPLE AND EFFECTIVE ENTITY ALIGNMENT FRAMEWORK FOR HETEROGENEOUS KNOWLEDGE GRAPHS
9804	AUDIO DEEPFAKE DETECTION WITH SELF-SUPERVISED WAVLM AND MULTI-FUSION ATTENTIVE CLASSIFIER
9821	AUDIO DIFFERENCE LEARNING FOR AUDIO CAPTIONING
4683	AUDIO MATCH CUTTING: FINDING AND CREATING MATCHING AUDIO TRANSITIONS IN MOVIES AND VIDEOS
9719	Audio prompt tuning for universal sound separation
7378	AUDIO TRANSFORMER FOR SYNTHETIC SPEECH DETECTION VIA FORMANT MAGNITUDE AND PHASE

<b>Paper Number</b>	<b>Paper Title</b>
	ANALYSIS
1114	AUDIO-AIDED LEARNING FRAMEWORK FOR IMAGE CLASSIFICATION WITH LIMITED TRAINING IMAGES
3101	AUDIO-FREE PROMPT TUNING FOR LANGUAGE-AUDIO MODELS
7774	Audio-Journey: Open Domain Latent Diffusion Based Text-to-Audio Generation
7441	AudioSR: Versatile Audio Super-resolution at Scale
3782	AUDIO-VISUAL ACTIVE SPEAKER EXTRACTION FOR SPARSELY OVERLAPPED MULTI-TALKER SPEECH
4865	AUDIO-VISUAL CHILD-ADULT SPEAKER CLASSIFICATION IN DYADIC INTERACTIONS
7799	AUDIOVISUAL SPEAKER SEPARATION WITH FULL- AND SUB-BAND MODELING IN THE TIME-FREQUENCY DOMAIN
7075	AUDIO-VISUAL SPEECH RECOGNITION IN-THE-WILD: MULTI-ANGLE VEHICLE CABIN CORPUS AND ATTENTION-BASED METHOD
6243	AUDITORY CORTEX-INSPIRED SPECTRAL ATTENTION MODULATION FOR BINAURAL SOUND LOCALIZATION IN HRTF MISMATCH
8024	Augment on Manifold: Mixup Regularization with UMAP
8356	AUGMENTING CONFORMERS WITH STRUCTURED STATE-SPACE SEQUENCE MODELS FOR ONLINE SPEECH RECOGNITION
7650	AUGMENTING TRANSFORMER AUTOENCODERS WITH PHENOTYPE CLASSIFICATION FOR ROBUST DETECTION OF PSYCHOTIC RELAPSES
7949	AUGSUMM: TOWARDS GENERALIZABLE SPEECH SUMMARIZATION USING SYNTHETIC LABELS FROM LARGE LANGUAGE MODELS
6396	AUTOCALI: ENHANCING AOA-BASED INDOOR LOCALIZATION THROUGH AUTOMATIC PHASE CALIBRATION
1102	AUTOFGNN: A FRAMEWORK FOR EXTRACTING ALL FREQUENCY INFORMATION FROM LARGE-SCALE GRAPHS

<b>Paper Number</b>	<b>Paper Title</b>
9686	Automated Labeling of Automotive Radar Azimuth Multipath
6254	AUTOMATIC CHANNEL SELECTION AND SPATIAL FEATURE INTEGRATION FOR MULTI-CHANNEL SPEECH RECOGNITION ACROSS VARIOUS ARRAY TOPOLOGIES
9249	AUTOMATIC DESIGN OF ADAPTER ARCHITECTURES FOR ENHANCED PARAMETER-EFFICIENT FINE-TUNING
3517	AUTOMATIC DETECTION OF SLEEPINESS-RELATED SYNDROMES AND SYMPTOMS USING VOICE AND SPEECH BIOMARKERS
2278	Automatic Recognition of Gesture Identity and Onset of Cued-Speech
8519	AUTOMATIC SPEECH RECOGNITION TUNED FOR CHILD SPEECH IN THE CLASSROOM
1434	AUTOMATIC TEMPORAL ALIGNMENT FOR PITCH ESTIMATION EVALUATION
8270	AUTOMOTIVE RADAR INTERFERENCE CHARACTERIZATION: FMCW OR PMCW?
8271	AUTOMOTIVE RADAR INTERFERENCE MITIGATION VIA SINR MAXIMIZATION
5615	AUTOMOTIVE RADAR POINT CLOUD PARAMETRIC DENSITY ESTIMATION USING CAMERA IMAGES
3290	AUTONOMOUS GENERATIVE FEATURE REPLAY FOR NON-EXEMPLAR CLASS-INCREMENTAL LEARNING
7570	AutoPrep: An Automatic Preprocessing Framework for In-the-Wild Speech Data
4358	AUTOREGRESSIVE 3D SHAPE COMPLETION VIA SPHERE-GUIDED DISENTANGLED REPRESENTATION
7667	AUTOSEN: IMPROVING AUTOMATIC WIFI HUMAN SENSING THROUGH CROSS-MODAL AUTOENCODER
8564	AutoSGM: A Unified Lowpass Regularization Framework for Accelerated Learning
4665	AutoST: Training-free Neural Architecture Search for Spiking Transformers
4347	AV2WAV: DIFFUSION-BASED RE-SYNTHESIS FROM CONTINUOUS SELF-SUPERVISED FEATURES FOR AUDIO-

<b>Paper Number</b>	<b>Paper Title</b>
	VISUAL SPEECH ENHANCEMENT
7520	AV-SUPERB: A MULTI-TASK EVALUATION BENCHMARK FOR AUDIO-VISUAL REPRESENTATION MODELS
5450	AXIS ORDER INVARIANCE LEARNED FROM POINT CLOUDS
3670	BAE-Net: A Low complexity and high fidelity bandwidth-adaptive neural network for speech super-resolution
6499	BALANCED AND DISCRIMINATIVE CONTRASTIVE LEARNING FOR CLASS-IMBALANCED MEDICAL IMAGES
9104	Balanced Learning for Multi-Domain Long-tailed Speaker Recognition
8159	Balancing Easy and Hard Distortions: A Multi-Rate Knowledge Distillation Strategy for Blind Image Quality Assessment
9743	Balancing Representation Abstractions and Local Details Preservation for 3D Point Cloud Quality Assessment
7478	Balancing Speaker-Rater Fairness for Gender-Neutral Speech Emotion Recognition
8020	BALLISTOCARDIOGRAM-BASED HEART RATE VARIABILITY ESTIMATION FOR STRESS MONITORING USING CONSUMER EARBUDS
2230	BANDWIDTH-EFFICIENT INFERENCE FOR NERUAL IMAGE COMPRESSION
7668	BASS ACCOMPANIMENT GENERATION VIA LATENT DIFFUSION
7876	BATCH SUBSTITUTION CALIBRATION OF A MEMS MICROPHONE ARRAY : IMPACT OF SENSOR PERFORMANCE DISPERSION ON DIRECTIVITY ESTIMATION
1703	Bayesian Activity Detection for Massive Connectivity in Cell-Free IoT Networks
8617	BAYESIAN LEARNING-BASED KALMAN SMOOTHING FOR LINEAR DYNAMICAL SYSTEMS WITH UNKNOWN SPARSE INPUTS
4049	BAYESIAN OPTIMIZATION WITH GAUSSIAN PROCESSES FOR ROBUST LOCALIZATION
11478	Bayesian Tensor Tucker Completion With a Flexible Core



<b>Paper Number</b>	<b>Paper Title</b>
7710	BAYESIAN TOPOLOGY INFERENCE ON PARTIALLY KNOWN NETWORKS FROM INPUT-OUTPUT PAIRS
4291	BAYESIAN-BOOSTED METALOC: EFFICIENT TRAINING AND GUARANTEED GENERALIZATION FOR INDOOR LOCALIZATION
5689	BCC: BIDIRECTIONAL CONSISTENCY CONSTRAINT METHOD FOR HIERARCHICAL TEXT CLASSIFICATION
6501	Beamforming Design and Performance Evaluation for RIS-aided Localization using LEO Satellite Signals
3326	Beamforming Through Online Convex Combination of Differential Beamformers
11480	BeamSync: Over-The-Air Synchronization for Distributed Massive MIMO Systems
2485	BEAST: ONLINE JOINT BEAT AND DOWNBEAT TRACKING BASED ON STREAMING TRANSFORMER
5332	BENCHMARKING ADVERSARIAL ROBUSTNESS OF IMAGE SHADOW REMOVAL WITH SHADOW-ADAPTIVE ATTACKS
9173	BETA QUANTILE REGRESSION FOR ROBUST ESTIMATION OF UNCERTAINTY IN THE PRESENCE OF OUTLIERS
6244	BEVLOC: END-TO-END 6-DOF LOCALIZATION VIA CROSS-MODALITY CORRELATION UNDER BIRD'S EYE VIEW
4166	BEVOXSEG: BEV-VOXEL REPRESENTATION FOR FAST AND ACCURATE CAMERA-BASED 3D SEGMENTATION
2937	BEYOND EMPIRICAL WINDOWING: AN ATTENTION-BASED APPROACH FOR TRUST PREDICTION IN AUTONOMOUS VEHICLES
7426	BEYOND SIMPLE TEXT STYLE TRANSFER: UNVEILING COMPOUND TEXT STYLE TRANSFER WITH PROMPT-BASED PRE-TRAINED LANGUAGE MODELS
4769	Beyond the Limit of Weight-Sharing: Pioneering Space-Evolving NAS with Large Language Models
3954	BEYOND THE SNOWFALL: ENHANCING SNOWY DAY OBJECT DETECTION THROUGH PROGRESSIVE RESTORATION AND MULTI-FEATURE FUSION.
6126	BFRFormer: Transformer-based generator for Real-World Blind Face Restoration

<b>Paper Number</b>	<b>Paper Title</b>
2607	BI-DIRECTIONAL MOTION ATTENTION WITH CONTRASTIVE LEARNING FOR FEW-SHOT ACTION RECOGNITION
1892	BigVSAN: Enhancing GAN-based Neural Vocoders with Slicing Adversarial Network
4511	BINARY SIGNAL ALIGNMENT: OPTIMAL SOLUTION IS POLYNOMIAL-TIME AND LINEAR-TIME SOLUTION IS QUASI-OPTIMAL
7856	Binaural Angular Separation Network
2969	BINAURAL RENDERING OF HETEROGENEOUS SOUND SOURCES WITH EXTENT
3955	BINAURAL ROOM TRANSFER FUNCTION INTERPOLATION VIA SYSTEM INVERSION
9962	BINAURAL SOUND SOURCE LOCALIZATION USING A HYBRID TIME AND FREQUENCY DOMAIN MODEL
4176	BINAURAL SPEECH ENHANCEMENT USING DEEP COMPLEX CONVOLUTIONAL TRANSFORMER NETWORKS
3697	BINAURALMUSIC: A DIVERSE DATASET FOR IMPROVING CROSS-MODAL BINAURAL AUDIO GENERATION
4910	Biomimetic Mappings for Active Sonar Object Recognition in Clutter
8605	BLENDa: DOMAIN ADAPTIVE OBJECT DETECTION THROUGH DIFFUSION-BASED BLENDING
3775	BLIND BEAMFORMING FOR INTELLIGENT REFLECTING SURFACE: A REINFORCEMENT LEARNING APPROACH
2831	BLIND DECONVOLUTION OF SPARSE GRAPH SIGNALS IN THE PRESENCE OF PERTURBATIONS
5690	Blind Estimation of Audio Effects using an Auto-Encoder Approach and Differentiable Digital Signal Processing
1234	BLIND INPAINTING WITH OBJECT-AWARE DISCRIMINATION FOR ARTIFICIAL MARKER REMOVAL
7299	BLIND SEPARATION OF NOISY MIXTURES OVER GALOIS FIELDS
5721	BLOCK ADAPTIVE SUBSPACE PURSUIT METHOD FOR WALL CLUTTER MITIGATION

<b>Paper Number</b>	<b>Paper Title</b>
11941	BMMSNet: Bidirectional Mapping and Multilevel Similarity Comparison for EEG-Speech Match-Mismatch Problem
7226	BNMTRANS: A BRAIN NETWORK SEQUENCE-DRIVEN MANIFOLD-BASED TRANSFORMER FOR COGNITIVE IMPAIRMENT DETECTION USING EEG
6381	BOOSTING ADVERSARIAL ROBUSTNESS DISTILLATION VIA HYBRID DECOMPOSED KNOWLEDGE
7509	Boosting End-to-End Multilingual Phoneme Recognition through Exploiting Universal Speech Attributes Constraints
7702	BOOSTING IMAGE QUALITY ASSESSMENT PERFORMANCE: UNSUPERVISED SCORE FUSION BY DEEP MAXIMUM A POSTERIORI ESTIMATION
8754	BOOSTING LLMS WITH ONTOLOGY-AWARE PROMPT FOR NER DATA AUGMENTATION
9505	Boosting of Implicit Neural Representation-based Image Denoiser
1365	BOOSTING PRUNED NETWORKS WITH LINEAR OVER-PARAMETERIZATION
8857	BOOSTING SPEECH ENHANCEMENT WITH CLEAN SELF-SUPERVISED FEATURES VIA CONDITIONAL VARIATIONAL AUTOENCODERS
2909	BOOSTING UNKNOWN-NUMBER SPEAKER SEPARATION WITH TRANSFORMER DECODER-BASED ATTRACTOR
5216	BOOSTING ZERO-SHOT HUMAN-OBJECT INTERACTION DETECTION WITH VISION-LANGUAGE TRANSFER
8827	BOOSTING ZERO-SHOT NODE CLASSIFICATION VIA DEPENDENCY CAPTURE AND DISCRIMINATIVE FEATURE LEARNING
6945	BOOTSTRAP PREDICTIVE CODING: INVESTIGATING A NON-CONTRASTIVE SELF-SUPERVISED LEARNING APPROACH
4453	BOUNDARY-DRIVEN ACTIVE LEARNING FOR ANOMALY DETECTION IN TIME SERIES DATA STREAMS
8910	BOUNDING BOX-GUIDED PSEUDO POINT CLOUDS EARLY-FUSION AND DENSITY OPTIMIZE FOR 3D OBJECT DETECTION

<b>Paper Number</b>	<b>Paper Title</b>
2228	BPDO:Boundary Points Dynamic Optimization for Arbitrary Shape Scene Text Detection
3201	BRAIN STRUCTURE-FUNCTION INTERACTION NETWORK FOR FLUID COGNITION PREDICTION
1220	BrainFC-CGAN: A Conditional Generative Adversarial Network for Brain Functional Connectivity Augmentation and Aging Synthesis
4712	BRANCHFORMER-BASED TDNN FOR AUTOMATIC SPEAKER VERIFICATION
6511	BRAVEN: IMPROVING SELF-SUPERVISED PRE-TRAINING FOR VISUAL AND AUDITORY SPEECH RECOGNITION
1007	Breaking Speaker Recognition with PaddingBack
1265	Breaking the Barrier: Selective Uncertainty-based Active Learning for Medical Image Segmentation
4587	BREAST ULTRASOUND COMPUTER-AIDED DIAGNOSIS USING STRUCTURE-AWARE TRIPLET PATH NETWORKS
4881	Bregman Graph Neural Network
8268	BRIDGING THE DOMAIN GAP ARISING FROM TEXT DESCRIPTION DIFFERENCES FOR STABLE TEXT-TO-IMAGE GENERATION
8516	BRIDGING THE GAP: A SELF-LEARNING MODEL USING IMPLICIT KNOWLEDGE FOR CHINESE SPELLING CORRECTION
9594	Bridging the Gap: Sketch to Color Diffusion Model with Semantic Prompt Learning
8238	BRIDGING THE GAPS OF BOTH MODALITY AND LANGUAGE: SYNCHRONOUS BILINGUAL CTC FOR SPEECH TRANSLATION AND SPEECH RECOGNITION
2409	BRINGING THE DISCUSSION OF MINIMA SHARPNESS TO THE AUDIO DOMAIN: A FILTER-NORMALISED EVALUATION FOR ACOUSTIC SCENE CLASSIFICATION
8947	Broadband Personal Sound Zone Control in the Presence of Nonlinearities
11872	BS-PLCNET: BAND-SPLIT PACKET LOSS CONCEALMENT NETWORK WITH MULTI-TASK LEARNING FRAMEWORK AND MULTI-DISCRIMINATORS

<b>Paper Number</b>	<b>Paper Title</b>
7904	Buffered Gaussian Modeling for Vectorized HD Map Construction
2179	BUILD A 50+ HOURS CHINESE MANDARIN CORPUS FOR CHILDREN'S SPEECH RECOGNITION
7377	Building Lane-Level Maps from Aerial Images
11916	BUMBLEBEE YOUR WAY TO RECOVERY: TRANSFORMING THE APPROACH TO DETECTION OF MENTAL HEALTH RELAPSES
2982	BWSNET: AUTOMATIC PERCEPTUAL ASSESSMENT OF AUDIO SIGNALS
7549	BYTEHUM: FAST AND ACCURATE QUERY-BY-HUMMING IN THE WILD
3043	CAGEN: CONTROLLABLE ANOMALY GENERATOR USING DIFFUSION MODEL
6682	CAG-FPN: CHANNEL SELF-ATTENTION GUIDED FEATURE PYRAMID NETWORK FOR OBJECT DETECTION
1826	CALSeg: Improving Calibration of Medical Image Segmentation Via Variational Label Smoothing
1515	CAMERA CALIBRATION USING A SINGLE VIEW OF A SYMMETRIC OBJECT
5217	CAMERA-RADAR ASSOCIATION FOR DATA ANNOTATION
2399	CAN CHATGPT SERVE AS A MULTI-CRITERIA DECISION MAKER? A NOVEL APPROACH TO SUPPLIER EVALUATION
2517	CAN LARGE-SCALE VOCODED SPOOFED DATA IMPROVE SPEECH SPOOFING COUNTERMEASURE WITH A SELF-SUPERVISED FRONT END?
7798	CAN LLM FIND THE GREEN CIRCLE? INVESTIGATION AND HUMAN-GUIDED TOOL MANIPULATION FOR COMPOSITIONAL GENERALIZATION
3975	CAN SYNTHETIC DATA BOOST THE TRAINING OF DEEP ACOUSTIC VEHICLE COUNTING NETWORKS?
2475	CAN WE TRUST EXPLAINABLE AI METHODS ON ASR? AN EVALUATION ON PHONEME RECOGNITION
8435	Can Whisper perform speech-based in-context learning?

<b>Paper Number</b>	<b>Paper Title</b>
4852	CAPTION UNIFICATION FOR MULTI-VIEW LIFELOGGING IMAGES BASED ON IN-CONTEXT LEARNING WITH HETEROGENEOUS SEMANTIC CONTENTS
7660	CAPTURING DETAIL VARIATIONS FOR LIGHTWEIGHT NEURAL RADIANCE FIELDS
6880	CARDINALITY-CONSTRAINED BINARY QUADRATIC OPTIMIZATION VIA EXTREME POINT PURSUIT, WITH APPLICATION TO THE DENSEST K-SUBGRAPH PROBLEM
7105	CARTOONDIFF: TRAINING-FREE CARTOON IMAGE GENERATION WITH DIFFUSION TRANSFORMER MODELS
9726	CAUSALITY-INSPIRED SINGLE-SOURCE DOMAIN GENERALIZATION FOR FACE ANTI-SPOOFING
3380	CAUSALLY UNCOVERING BIAS IN VIDEO MICRO-EXPRESSION RECOGNITION
3383	CAUSALME: BALANCING BI-MODALITIES IN VISUAL QUESTION ANSWERING
4174	CAUSAL-STORY: LOCAL CAUSAL ATTENTION UTILIZING PARAMETER-EFFICIENT TUNING FOR VISUAL STORY SYNTHESIS
6080	CC-DA: CROSS-DOMAIN CONSISTENCY DATA AUGMENTATION FOR 3D TUMOR SEGMENTATION
4167	C-CLAPA: IMPROVING TEXT-AUDIO CROSS DOMAIN RETRIEVAL WITH CAPTIONING AND AUGMENTATIONS
6238	CDA-MBPO:CORRECTED DATA AGGREGATION FOR MODEL-BASED POLICY OPTIMIZATION
2693	CDCNet: A FAST and LIGHTWEIGHT DEHAZING NETWORK WITH COLOR DISTORTION CORRECTION
3394	CDUMA: An Adaptive Approach for Mitigating Confounder for MCQA
1532	CED: Consistent ensemble distillation for audio tagging
7050	CEDNET: A CONTINUOUS EMOTION DETECTION NETWORK FOR NATURALISTIC STIMULI USING MEG SIGNALS
4685	CEMOAE: A DYNAMIC AUTOENCODER WITH MASKED CHANNEL MODELING FOR ROBUST EEG-BASED EMOTION RECOGNITION

<b>Paper Number</b>	<b>Paper Title</b>
10015	CENET: CONTENT-AWARE ENHANCED NETWORK FOR PRACTICAL SCENE PARSING
4669	CENTER OF PRESSURE ESTIMATION BY ANALYZING WALKING VIDEOS
8921	CGN: A SIMPLE YET EFFECTIVE MULTI-CHANNEL GATED NETWORK FOR LONG-TERM TIME SERIES FORECASTING
1573	CHANGENET: MULTI-TEMPORAL ASYMMETRIC CHANGE DETECTION DATASET
4339	CHANNEL ESTIMATION AND PREDICTION IN WIRELESS COMMUNICATIONS ASSISTED BY SEMI-PASSIVE RIS
5652	CHANNEL ESTIMATION IN UNDERDETERMINED SYSTEMS UTILIZING VARIATIONAL AUTOENCODERS
1446	CHANNEL-SPATIAL TRANSFORMER FOR EFFICIENT IMAGE SUPER-RESOLUTION
8171	Character Attribute Extraction from Movie Scripts using LLMs
2199	CHAT: Cascade Hole-Aware Transformers with Geometric Spatial Consistency for Accurate Monocular Endoscopic Depth Estimation
1647	CHILD FER: DOMAIN-AGNOSTIC FACIAL EXPRESSION RECOGNITION IN CHILDREN USING A SECONDARY IMAGE DIFFUSION MODEL
5999	CHUNKED ATTENTION-BASED ENCODER-DECODER MODEL FOR STREAMING SPEECH RECOGNITION
3388	CIF-RNNT: Streaming ASR via Acoustic Word Embeddings with Continuous Integrate-and-Fire and RNN-Transducers
3303	CIF-T: A Novel CIF-based Transducer Architecture for Automatic Speech Recognition
5214	CKT-RCM: CLIP-BASED KNOWLEDGE TRANSFER AND RELATIONAL CONTEXT MINING FOR UNBIASED PANOPTIC SCENE GRAPH GENERATION
8317	CLAF: CONTRASTIVE LEARNING WITH AUGMENTED FEATURES FOR IMBALANCED SEMI-SUPERVISED LEARNING
7363	CLAP4EMO: CHATGPT-ASSISTED SPEECH EMOTION RETRIEVAL WITH NATURAL LANGUAGE SUPERVISION

<b>Paper Number</b>	<b>Paper Title</b>
9224	CLASS: CONTINUAL LEARNING APPROACH FOR SPEECH SUPER-RESOLUTION
7413	CLASSIFICATION-ORIENTED SEMANTIC WIRELESS COMMUNICATIONS
6170	CLASS-INCREMENTAL LEARNING FOR MULTI-LABEL AUDIO CLASSIFICATION
7260	CLASS-WISE BUFFER MANAGEMENT FOR INCREMENTAL OBJECT DETECTION: AN EFFECTIVE BUFFER TRAINING STRATEGY
10279	CLIENT-FREE FEDERATED UNLEARNING VIA TRAINING RECONSTRUCTION WITH ANCHOR SUBSPACE CALIBRATION
7315	CLINICAL SCORES PREDICTION AND MEDICATION ADJUSTMENT FOR COURSE OF PARKINSON'S DISEASE
2620	CLIP-BASED SYNERGISTIC KNOWLEDGE TRANSFER FOR TEXT-BASED PERSON RETRIEVAL
5672	CLIP-FONT: SEMANTIC SELF-SUPERVISED FEW-SHOT FONT GENERATION WITH CLIP
5963	CLIP-MSA: INCORPORATING INTER-MODAL DYNAMICS AND COMMON KNOWLEDGE TO MULTIMODAL SENTIMENT ANALYSIS WITH CLIP
1386	CLIPRerank: An Extremely Simple Method for Improving Ad-hoc Video Search
11558	Closed-Loop Training for Projected GAN
6566	CLOSE-RANGE DIRECTION OF ARRIVAL ESTIMATION IN THE PRESENCE OF CLOCK JITTER
6201	CLPSD: DETECTING ETHEREUM PHISHING SCAMS BASED ON CURRICULUM LEARNING
1443	CLT: COOPERATIVE LOTTERY TICKET HYPOTHESIS IN LIVE STREAMING SALES PREDICTION
10446	CLUSTER-GUIDED UNSUPERVISED DOMAIN ADAPTATION FOR DEEP SPEAKER EMBEDDING
9858	CM-PIE: CROSS-MODAL PERCEPTION FOR INTERACTIVE-ENHANCED AUDIO-VISUAL VIDEO PARSING
3816	CNFA: Conditional Normalizing Flow for Query-Limited Attack



<b>Paper Number</b>	<b>Paper Title</b>
9306	CODING FOR THE UNSOURCED B-CHANNEL WITH ERASURES: ENHANCING THE LINKED LOOP CODE
7009	COGNITIVE VIRTUAL SENSING TECHNIQUE FOR FEEDFORWARD ACTIVE NOISE CONTROL
5509	COLLABORATIVE WATERMARKING FOR ADVERSARIAL SPEECH SYNTHESIS
4330	COLLD: CONTRASTIVE LAYER-TO-LAYER DISTILLATION FOR COMPRESSING MULTILINGUAL PRE-TRAINED SPEECH ENCODERS
5219	COLOR AGNOSTIC CROSS-SPECTRAL DISPARITY ESTIMATION
1599	ColorFlow: A Conditional Normalizing Flow for Image Colorization
6836	Combining Conformer and Dual-Path-Transformer Networks for Single Channel Noisy Reverberant Speech Separation
6903	COMMIN: SEMANTIC IMAGE COMMUNICATIONS AS AN INVERSE PROBLEM WITH INN-GUIDED DIFFUSION MODELS
11484	COMMON-SLOPE MODELING OF LATE REVERBERATION
9520	Communication Efficient Private Federated Learning Using Dithering
8205	COMMUNICATION-EFFICIENT DECENTRALIZED DYNAMIC KERNEL LEARNING
3451	COMMUNICATION-EFFICIENT FEDERATED LEARNING THROUGH ADAPTIVE WEIGHT CLUSTERING AND SERVER-SIDE DISTILLATION
7479	Communication-Efficient Federated Optimization over Semi-Decentralized Networks
3060	COMMUNICATION-EFFICIENT LAPLACE MECHANISM FOR DIFFERENTIAL PRIVACY VIA RANDOM QUANTIZATION
1416	Communication-Efficient Personalized Federated Learning for Speech-to-Text Tasks
7564	COMMUNICATION-ORIENTED AUTOMATIC ASSESSMENT SYSTEM FOR ACCENTED SPOKEN CHINESE IN READ-ALoud TASKS

<b>Paper Number</b>	<b>Paper Title</b>
10226	COMPACT AND DE-BIASED NEGATIVE INSTANCE EMBEDDING FOR MULTI-INSTANCE LEARNING ON WHOLE-SLIDE IMAGE CLASSIFICATION
3044	COMPARABLE DEMONSTRATIONS ARE IMPORTANT IN IN-CONTEXT LEARNING: A NOVEL PERSPECTIVE ON DEMONSTRATION SELECTION
8923	COMPARATIVE STUDY OF TOKENIZATION ALGORITHMS FOR END-TO-END OPEN VOCABULARY KEYWORD DETECTION
8418	COMPARING AND COMBINING AUDIO PROCESSING AND DEEP LEARNING FEATURES FOR CLASSIFICATION OF HEARTBEAT SOUNDS
7423	Comparing data-driven and handcrafted features for dimensional emotion recognition
7402	COMPARISON OF CONDITIONS FOR OMNIDIRECTIONAL VIDEO WITH SPATIAL AUDIO IN TERMS OF SUBJECTIVE QUALITY AND IMPACTS ON OBJECTIVE METRICS RESOLVING POWER
4335	COMPARISON OF FREQUENCY-FUSION MECHANISMS FOR BINAURAL DIRECTION-OF-ARRIVAL ESTIMATION FOR MULTIPLE SPEAKERS
1392	Complementary Fusion Network based on Frequency Hybrid Attention for Pansharpening
6905	Complex Bounded Component Analysis: Identifiability and Algorithm
3795	COMPLEXITY REDUCTION OF TEMPLATE MATCHING-BASED REFERENCE PICTURE PADDING IN VIDEO CODING
8497	Complexity Scaling for Speech Denoising
3521	COMPOSITE FEDERATED LEARNING WITH HETEROGENEOUS DATA
11544	COMPRESSION OF HIGHER-ORDER AMBISONIC SIGNALS USING DIRECTIONAL AUDIO CODING
9531	COMPUTATIONAL COMPLEXITY OF ASYNCHRONOUS POLICY ITERATION FOR TWO-PLAYER ZERO-SUM MARKOV GAMES
3382	COMPUTING AN ENTIRE SOLUTION PATH OF A

<b>Paper Number</b>	<b>Paper Title</b>
	NONCONVEXLY REGULARIZED CONVEX SPARSE MODEL
9035	CONCEALING MEDICAL CONDITION BY NODE TOGGING IN ASR FOR DEMENTIA PATIENTS
6848	CONCENTRATED REASONING AND UNIFIED RECONSTRUCTION FOR MULTI-MODAL MEDIA MANIPULATION
4017	CONCSS: CONTRASTIVE-BASED CONTEXT COMPREHENSION FOR DIALOGUE-APPROPRIATE PROSODY IN CONVERSATIONAL SPEECH SYNTHESIS
3933	CONFIDENCE-AWARE SPATIAL-TEMPORAL ATTENTION GRAPH CONVOLUTIONAL NETWORK FOR SKELETON-BASED EXPERT-NOVICE LEVEL CLASSIFICATION
7850	CONFORMALIZED MULTIMODAL UNCERTAINTY REGRESSION AND REASONING
1995	Conformer is all you need for visual speech recognition
3624	CONGESTION-AWARE DISTRIBUTED TASK OFFLOADING IN WIRELESS MULTI-HOP NETWORKS USING GRAPH NEURAL NETWORKS
3545	Conjugate Gradient Based Adaptive Algorithm for Nonlinear AEC
9578	CONNECTING SPEECH ENCODER AND LARGE LANGUAGE MODEL FOR ASR
6467	CONSIDERING TEMPORAL CONNECTION BETWEEN TURNS FOR CONVERSATIONAL SPEECH SYNTHESIS
6859	CONSISTENT AND RELEVANT: RETHINK THE QUERY EMBEDDING IN GENERAL SOUND SEPARATION
7366	ConsPrompt: Exploiting Contrastive Samples for Few-shot Prompt Learning
6009	CONTACTLESS RADAR HEART RATE VARIABILITY MONITORING VIA DEEP SPATIO-TEMPORAL MODELING
7078	CONTENT-BASED OBJECTIVE EVALUATION OF ARTIFICIALLY GENERATED SIGN LANGUAGE VIDEOS
7148	CONTEXT-AWARE AND CONTRASTIVENESS-DRIVEN FEATURE LEARNING FOR CROSS-DOMAIN FEW-SHOT HYPERSPECTRAL IMAGE CLASSIFICATION

<b>Paper Number</b>	<b>Paper Title</b>
10169	CONTEXT-AWARE DUAL ATTENTION NETWORK FOR MULTIMODAL SARCASM DETECTION
7268	CONTEXT-AWARE PREFERENCE LEARNING SYSTEM BASED ON DIRICHLET PROCESS GAUSSIAN MIXTURE MODEL
9727	CONTEXT-AWARE TRANSFORMER FOR SINGLE IMAGE RAIN STREAKS REMOVAL
8862	CONTEXT-GUIDED AND SYNTACTIC AUGMENTED DUAL GRAPH CONVOLUTIONAL NETWORK FOR ASPECT-BASED SENTIMENT ANALYSIS
9617	CONTEXTUAL BIASING METHODS FOR IMPROVING RARE WORD DETECTION IN AUTOMATIC SPEECH RECOGNITION
2008	Contextual Biasing of Named-Entities with Large Language Models
8630	CONTEXTUAL HUMAN OBJECT INTERACTION UNDERSTANDING FROM PRE-TRAINED LARGE LANGUAGE MODEL
4546	CONTEXTUALIZED AUTOMATIC SPEECH RECOGNITION WITH ATTENTION-BASED BIAS PHRASE BOOSTED BEAM SEARCH
9098	CONTINUAL LEARNING WITH CLASS-LEVEL MINIMALLY INTERFERED UPDATE
3161	Continuous Review and Timely Correction: Enhancing the Resistance to Noisy Labels via Self-Not-True Distillation
7027	CONTRASTIVE DEEP NONNEGATIVE MATRIX FACTORIZATION FOR COMMUNITY DETECTION
1403	CONTRASTIVE LEARNING FOR REGRESSION ON HYPERSPECTRAL DATA
7146	CONTRASTIVE LEARNING WITH AUDIO DISCRIMINATION FOR CUSTOMIZABLE KEYWORD SPOTTING IN CONTINUOUS SPEECH
1295	CONTRASTIVE LEARNING WITH BIDIRECTIONAL TRANSFORMERS FOR KNOWLEDGE TRACING
6941	CONTRASTIVE LEARNING WITH HIGH-QUALITY AND LOW-QUALITY AUGMENTED DATA FOR QUERY-FOCUSED SUMMARIZATION

<b>Paper Number</b>	<b>Paper Title</b>
7145	CONTRASTIVE LOSS BASED FRAME-WISE FEATURE DISENTANGLEMENT FOR POLYPHONIC SOUND EVENT DETECTION
4544	CONTRASTIVE SPEAKER EMBEDDING WITH SEQUENTIAL DISENTANGLEMENT
9258	CONTRMIX: PROGRESSIVE MIXED CONTRASTIVE LEARNING FOR SEMI-SUPERVISED MEDICAL IMAGE SEGMENTATION
8858	CONTROLCAP: CONTROLLABLE CAPTIONING VIA NO-FUSS LEXICON
7653	CONTROLLABLE PROSODY GENERATION WITH PARTIAL INPUTS
3149	CONTROLLABLE SEMANTIC LINGUISTIC STEGANOGRAPHY VIA SUMMARIZATION GENERATION
4411	CONTROLLABLE SPEAKING STYLES USING A LARGE LANGUAGE MODEL
11938	CONVCONCATNET: A DEEP CONVOLUTIONAL NEURAL NETWORK TO RECONSTRUCT MEL SPECTROGRAM FROM THE EEG
7594	CONVERGENT PLUG-AND-PLAY USING CONTRACTIVE DENOISERS
3710	CONVERSATION CLIQUE-BASED MODEL FOR EMOTION RECOGNITION IN CONVERSATION
8375	CONVERSATIONAL CO-SPEECH GESTURE GENERATION VIA MODELING DIALOG INTENTION, EMOTION AND CONTEXT WITH DIFFUSION MODELS
8956	CONVNEXT-TTS AND CONVNEXT-VC: CONVNEXT-BASED FAST END-TO-END SEQUENCE-TO-SEQUENCE TEXT-TO-SPEECH AND VOICE CONVERSION
11456	CONVOLUTIONAL FILTERS AND NEURAL NETWORKS WITH NONCOMMUTATIVE ALGEBRAS
1999	Co-occurrence Graph-Enhanced Hierarchical Prediction of ICD Codes
6616	COOKING-CLIP: CONTEXT-AWARE LANGUAGE-IMAGE PRETRAINING FOR ZERO-SHOT RECIPE GENERATION

<b>Paper Number</b>	<b>Paper Title</b>
1910	Cooperative Sensing via Matrix Factorization of the Partially Received Sample Covariance Matrix
9630	COORDINATE-BASED NEURAL NETWORK FOR FOURIER PHASE RETRIEVAL
2400	COPHTC: CONTRASTIVE LEARNING WITH PROMPT TUNING FOR HIERARCHICAL TEXT CLASSIFICATION
6982	COQ:AN EMPIRICAL FRAMEWORK FOR MULTI-HOP QUESTION ANSWERING EMPOWERED BY LARGE LANGUAGE MODELS
8053	CORAAL QA: A Dataset and Framework for Open Domain Spontaneous Speech Question Answering from Long Audio Files
1857	CORE BODY TEMPERATURE AND ITS ROLE IN DETECTING ACUTE STRESS: A FEASIBILITY STUDY
2156	CORN: CO-TRAINED FULL- AND NO-REFERENCE SPEECH QUALITY ASSESSMENT
4516	CORNER DETECTION BASED ON A ROTATION-INVARIANT AND NOISE-INSENSITIVE CURVATURE MEASUREMENT
8621	Corpus Synthesis for Zero-shot ASR Domain Adaptation using Large Language Models
1518	CORRECTING FAULTY ROAD MAPS BY IMAGE INPAINTING
4412	CORRECTION FOCUSED LANGUAGE MODEL TRAINING FOR SPEECH RECOGNITION
2005	CORRELATION-BASED MACHINE LEARNING TECHNIQUES FOR CHANNEL ESTIMATION WITH FLUID ANTENNAS
8905	CO-SALIENT OBJECT DETECTION VIA DISCRIMINATIVE PROTOTYPES CONTRAST
1329	CoSLR: Contrastive Chinese Sign Language Recognition with Prior Knowledge and Multi-tasks Joint Learning
10005	COST AWARE UNTARGETED POISONING ATTACK AGAINST GRAPH NEURAL NETWORKS
7977	Counting Network for Learning from Majority Label
7890	COUPLED BLOCK-TERM TENSOR DECOMPOSITION FOR NEAR-FIELD LOCALIZATION IN MULTI-STATIC MIMO RADAR SYSTEMS

<b>Paper Number</b>	<b>Paper Title</b>
9708	Coupling Self-Supervised and Supervised Contrastive Learning for Multiple Classification of Cervical Cytological Whole Slide Images
11492	Covariance Matrix Recovery From One-Bit Data With Non-Zero Quantization Thresholds: Algorithm and Performance Analysis
9099	COVERAGE ANALYSIS FOR MMWAVE UAV NETWORKS WITH STATIC AND DYNAMIC BLOCKAGES
11553	COVID-19 Detection via Fusion of Modulation Spectrum and Linear Prediction Speech Features
4729	CPAUG: REFINING COPY-PASTE AUGMENTATION FOR SPEECH ANTI-SPOOFING
10199	CPMSVD: Cross-Project Multiclass Software Vulnerability Detection via Fused Deep Feature and Domain Adaptation
9376	CRAMER-RAO BOUND FOR ADMITTANCE MATRIX ESTIMATION UNDER LAPLACIAN CONSTRAINTS
1124	CRC-AIDED LEARNED ENSEMBLES OF BELIEF-PROPAGATION POLAR DECODERS
6849	CREATING PERSONALIZED SYNTHETIC VOICES FROM ARTICULATION IMPAIRED SPEECH USING AUGMENTED RECONSTRUCTION LOSS
1582	Credible Teacher for Semi-Supervised Object Detection in Open Scene
8753	CRESTYLER: TEXT-GUIDED SINGLE IMAGE STYLE TRANSFER METHOD BASED ON CNN AND RESTORMER
3000	CroCFuN: Cross-modal Conditional Fusion Network for Pansharpening
1879	CROSS BRANCH FEATURE FUSION DECODER FOR CONSISTENCY REGULARIZATION-BASED SEMI-SUPERVISED CHANGE DETECTION
8411	Cross Modal Training For ASR Error Correction With Contrastive Learning
9143	Cross Pseudo-Labeling for Semi-Supervised Audio-Visual Source Localization
3621	CROSS-AGE CONTRASTIVE LEARNING FOR AGE-INVARIANT FACE RECOGNITION

<b>Paper Number</b>	<b>Paper Title</b>
3869	CROSS-ATTENTION WATERMARKING OF LARGE LANGUAGE MODELS
11858	CROSS-ATTENTION-GUIDED WAVENET FOR MEL SPECTROGRAM RECONSTRUCTION IN THE ICASSP 2024 AUDITORY EEG CHALLENGE
1121	Cross-Camera Human Motion Transfer by Time Series Analysis
1284	CROSS-DOMAIN CROSS-TASK TRANSFER MOBILE TOUCH-STROKE AUTHENTICATION
7143	CROSS-IMAGE DISTILLATION FOR SEMI-SUPERVISED SEMANTIC SEGMENTATION
7208	CROSS-LINGUAL LEARNING IN MULTILINGUAL SCENE TEXT RECOGNITION
11873	Cross-lingual Text-to-Speech via Hierarchical Style Transfer
3459	CROSS-MODAL ALIGNMENT FOR END-TO-END SPOKEN LANGUAGE UNDERSTANDING BASED ON MOMENTUM CONTRASTIVE LEARNING
9690	CROSS-MODAL MULTISCALE DIFFERENCE-AWARE NETWORK FOR JOINT MOMENT RETRIEVAL AND HIGHLIGHT DETECTION
7698	CROSS-MODAL MULTI-TASKING FOR SPEECH-TO-TEXT TRANSLATION VIA HARD PARAMETER SHARING
2881	CROSS-MODAL PARALLEL TRAINING FOR IMPROVING END-TO-END ACCENTED SPEECH RECOGNITION
3815	Cross-Modal Synthesis of Structural MRI and Functional Connectivity Networks via Conditional ViT-GANs
9369	CROSS-MODALITY AND WITHIN-MODALITY REGULARIZATION FOR AUDIO-VISUAL DEEPFAKE DETECTION
7789	Cross-speaker encoding network for multi-talker speech recognition
3731	CROSS-SUBJECT EEG EMOTION RECOGNITION BASED ON INTERCONNECTED DYNAMIC DOMAIN ADAPTATION
3703	CROSS-TARGET STANCE DETECTION BY EXPLOITING TARGET ANALYTICAL PERSPECTIVES



<b>Paper Number</b>	<b>Paper Title</b>
5935	CROSS-TRIGGERING ISSUE IN AUDIO EVENT DETECTION AND MITIGATION
8460	CROSSWORD: A SEMANTIC APPROACH TO TEXT COMPRESSION VIA MASKING
4207	CROWD MODELING AND CONTROL VIA COOPERATIVE ADAPTIVE FILTERING
8502	Crowdsourced and Automatic Speech Prominence Estimation
9588	Crowdsourced multilingual speech intelligibility testing
7740	CryCeleb: A Speaker Verification Dataset Based on Infant Cry Sounds
7845	CRYPTO-MINE: Cryptanalysis via Mutual Information Neural Estimation
6196	CSCNET: CLASS-SPECIFIED CASCADED NETWORK FOR COMPOSITIONAL ZERO-SHOT LEARNING
4392	CSI-Free Over-the-Air Decentralized Learning over Frequency Selective Channels
10370	CSNET: CONTRASTIVE SIAMESE NETWORK FOR ROBUST SLU
6838	CST-FORMER: TRANSFORMER WITH CHANNEL-SPECTRO-TEMPORAL ATTENTION FOR SOUND EVENT LOCALIZATION AND DETECTION
3276	CT AND MRI FUSION WITH ANISOTROPIC GUIDED FILTERING
3165	CUBIC KNOWLEDGE DISTILLATION FOR SPEECH EMOTION RECOGNITION
8457	CUFFLESS BLOOD PRESSURE ESTIMATION USING MAGNETIC FLUX IN A RING FORM FACTOR
3164	Curricular Contrastive Regularization for Speech Enhancement with Self-supervised Representations
5958	Customising General Large Language Models for Specialised Emotion Recognition Tasks
3196	Customized Treatment Per Pixel for Blind Image Super-Resolution

<b>Paper Number</b>	<b>Paper Title</b>
4148	CutDEM: Depth-Aware Enhanced Multi-View Image Mixing for Light Field Super-Resolution
8898	CUTransNet: Transformers to Make Strong Encoders for Multi-Task Vision Perception of Autonomous Driving
9809	CYCLIC MISSPECIFIED CRAMER-RAO BOUND FOR PERIODIC PARAMETER ESTIMATION
1781	D3: DUAL-DOMAIN DEFENSES FOR BYZANTINE-RESILIENT DECENTRALIZED RESOURCE ALLOCATION
9454	DACR: DISTRIBUTION-AUGMENTED CONTRASTIVE RECONSTRUCTION FOR TIME-SERIES ANOMALY DETECTION
2898	DAMP: DISTRIBUTION-AWARE MAGNITUDE PRUNING FOR BUDGET-SENSITIVE GRAPH CONVOLUTIONAL NETWORKS
1213	DAP: Domain-aware Prompt Learning for Vision-and-Language Navigation
2815	DarkShot: Lighting Dark Images with Low-Compute and High-Quality
8730	DATA AUGMENTATION VIA SUBGROUP MIXUP FOR IMPROVING FAIRNESS
4931	DATA DRIVEN GRAPHEME-TO-PHONEME REPRESENTATIONS FOR A LEXICON-FREE TEXT-TO-SPEECH
1947	DATA-AIDED CHANNEL ESTIMATION UTILIZING GAUSSIAN MIXTURE MODELS
8101	DATA-DRIVEN CONVEX REGULARIZERS FOR INVERSE PROBLEMS
4830	DATA-DRIVEN LATTICES FOR VECTOR QUANTIZATION
2132	DATA-FREE WATERMARK FOR DEEP NEURAL NETWORKS BY TRUNCATED ADVERSARIAL DISTILLATION
6934	DATA-SCARCE CONDITION MODELING REQUIRES MODEL-BASED PRIOR REGULARIZATION
2639	Dataset Distillation with Channel Efficient Process
3469	DBS: Differentiable Budget-aware Searching for channel pruning
4726	DCL-NET: DUAL CONTRASTIVE LEARNING NETWORK FOR SEMI-SUPERVISED MULTI-ORGAN SEGMENTATION

<b>Paper Number</b>	<b>Paper Title</b>
2962	DCS: DEBIASED CONTRASTIVE LEARNING WITH WEAK SUPERVISION FOR TIME SERIES CLASSIFICATION
6030	DCTTS: DISCRETE DIFFUSION MODEL WITH CONTRASTIVE LEARNING FOR TEXT-TO-SPEECH GENERATION
4823	DDD: A PERCEPTUALLY SUPERIOR LOW-RESPONSE-TIME DNN-BASED DECLIPPER
7562	DDI-COCO: A DATASET FOR UNDERSTANDING THE EFFECT OF COLOR CONTRAST IN MACHINE-ASSISTED SKIN DISEASE DETECTION
2356	DDN-Net: Deep Residual Shrinkage Denoising Networks with Channel-wise Adaptively Soft Thresholds for Automated Major Depressive Disorder Identification
8068	DE NOVO MOLECULE GENERATION WITH GRAPH LATENT DIFFUSION MODEL
4883	DEBIASING RECOMMENDERS THROUGH PERSONALIZED POPULARITY-AWARE MARGINS
6644	Debris sensing based on LEO constellation: an intersatellite channel parameter estimation approach
2044	DECENTRALIZED GENERALIZED APPROXIMATE MESSAGE-PASSING FOR TREE-STRUCTURED NETWORKS
3025	DECENTRALIZED LOW RANK MATRIX RECOVERY FROM COLUMN-WISE PROJECTIONS BY ALTERNATING GD AND MINIMIZATION
3325	DECENTRALIZING COHERENT JOINT TRANSMISSION PRECODING VIA DETERMINISTIC EQUIVALENTS
2422	DECOUPLED SELF-ADAPTIVE DISTRIBUTION REGULARIZATION FOR FEW-SHOT IMAGE CLASSIFICATION
8326	DECOUPLED SPATIAL AND TEMPORAL PROCESSING FOR RESOURCE EFFICIENT MULTICHANNEL SPEECH ENHANCEMENT
9380	Decoupling and Refilling: A Simple Data Augmentation Method for Aspect Term Extraction
3668	Deep convolution network based super resolution DOA estimation with Toeplitz and sparse prior
2776	DEEP FUSION OF SHIFTED MLP AND CNN FOR MEDICAL IMAGE SEGMENTATION

<b>Paper Number</b>	<b>Paper Title</b>
9301	DEEP INCM RECONSTRUCTION FOR ADAPTIVE BEAMFORMING
4341	DEEP LEARNING AMR MODEL INFERENCE ACCELERATION WITH CFU FOR EDGE SYSTEMS
8430	Deep learning based single-shot profilometry by three-channel binary-defocused projection
7371	DEEP LEARNING INVERSION OF OCEAN WAVE SPECTRUM FROM SAR SATELLITE OBSERVATIONS
4036	DEEP MANIFOLD TRANSFORMATION FOR PROTEIN REPRESENTATION LEARNING
10200	DEEP NEIGHBOR LAYER AGGREGATION FOR LIGHTWEIGHT SELF-SUPERVISED MONOCULAR DEPTH ESTIMATION
2127	DEEP NEURAL NETWORK MODELS TRAINED WITH A FIXED RANDOM CLASSIFIER TRANSFER BETTER ACROSS DOMAINS
6650	Deep Optimization of relay networks - Using Relays as Neurons
11457	DEEP ORDINAL REGRESSION FRAMEWORK FOR NO-REFERENCE IMAGE QUALITY ASSESSMENT
4337	Deep Plug-and-Play Algorithm for Unsaturated Imaging
9252	Deep regression for biological age estimation in multiple organs: Investigations on 40,000 subjects of the UK Biobank
4924	Deep Reinforcement Learning for Energy Minimization in Multi-RIS-Aided Cell-Free MEC Networks
3801	DEEP RESIDUAL W-UNIT LEARNING WITH SEMANTIC EMBEDDING FOR AUTOMATIC PULMONARY CT ARTERY-VEIN SEPARATION
7333	DEEP UNFOLDED ANNEALED STEIN PARTICLE FILTER FOR VEHICLE TRACKING
8633	DEEP UNROLLING NETWORK FOR SAR IMAGE DESPECKLING
9607	DEEP VARIATIONAL PRIVACY FUNNEL: GENERAL MODELING WITH APPLICATIONS IN FACE RECOGNITION
1717	Deep Versatile Hyperspectral Reconstruction Model from a Snapshot Measurement with Arbitrary Masks

<b>Paper Number</b>	<b>Paper Title</b>
11479	DEEPCOMBOSAD: SPECTRO-TEMPORAL CORRELATION BASED SPEECH ACTIVITY DETECTION FOR NATURALISTIC AUDIO STREAMS
7870	DEEPGRE: GLOBAL ROBUSTNESS EVALUATION OF DEEP NEURAL NETWORKS
2940	DEEPOREDNET: CONTRASTIVE LEARNING-BASED ATTENTION-WEIGHTED DUAL CHANNEL RESIDUAL NETWORK FOR OCULAR REDNESS ASSESSMENT
2610	DEFENDING AGAINST CLEAN-IMAGE BACKDOOR ATTACK IN MULTI-LABEL CLASSIFICATION
5855	DEFOCUSSR: An EFFICIENT FRAMEWORK FOR DEFOCUS IMAGE SUPER-RESOLUTION GUIDED BY DEPTH INFORMATION
3489	DEFORMATION AND PENETRATION HYBRID DETECTION-NET FOR PARCELS INSPECTION IN INDUSTRIAL SUPPLY CHAIN
1759	DEFORMMLP: DYNAMIC LARGE-SCALE RECEPTIVE FIELD MLP NETWORKS FOR HUMAN MOTION PREDICTION
8687	DEGAN: DISCRIMINATION ENHANCED GAN FOR PERCEPTUAL-ORIENTED SUPER-RESOLUTION
4648	DELAY EMBEDDING FOR MATRIX GRAPHICAL MODEL LEARNING FROM DEPENDENT DATA
11546	Delayless Generative Fixed-filter Active Noise Control based on Deep Learning and Bayesian Filter
9387	DELINEATION OF PROSTATE CANCER VIA ENHANCED AI-BASED ALGORITHM IN ULTRASOUND IMAGES
2252	DELVING DEEPER INTO VULNERABLE SAMPLES IN ADVERSARIAL TRAINING
8973	DEMENTIA ASSESSMENT USING MANDARIN SPEECH WITH AN ATTENTION-BASED SPEECH RECOGNITION ENCODER
11919	DEMUCS for Data-Driven RF Signal Denoising
9026	Denoising Diffusion Probabilistic Models for Action-Conditioned 3D Motion Generation
4691	Depth-guided dominant plane perception for unsupervised homography estimation

<b>Paper Number</b>	<b>Paper Title</b>
8779	DESIGN OF SPATIAL-SLOW-TIME CONSTANT-MODULUS WAVEFORM TRANSMISSION AND RECEIVE ADAPTIVE FILTER FOR DUAL-FUNCTION RADAR COMMUNICATIONS WITH RECONFIGURABLE INTELLIGENT SURFACE
7808	DETECTING CHECK-WORTHY CLAIMS IN POLITICAL DEBATES, SPEECHES, AND INTERVIEWS USING AUDIO DATA
7687	DETECTING CONTINUOUS GRAVITATIONAL WAVES USING GENERATED TRAINING DATA
11892	DETECTING GAMMA-BAND RESPONSES TO THE SPEECH ENVELOPE FOR THE ICASSP 2024 AUDITORY EEG DECODING SIGNAL PROCESSING GRAND CHALLENGE
9050	DETECTION AND ATTRIBUTION OF MODELS TRAINED ON GENERATED DATA
6079	DETECTION IN COMPLEX SCENES USING RGB AND DEPTH MULTIMODAL FEATURE FUSION
9002	DETECTION OF EPILEPTIC SEIZURES IN LONG EEG RECORDINGS USING AN ANOMALY DETECTOR WITH ARTIFACT REJECTION
2879	DETECTOR DESIGN FOR DISTRIBUTED MULTICHANNEL RADAR SENSORS IN COLORED INTERFERENCE ENVIRONMENTS
5842	DETERMINED BSS BY COMBINATION OF IVA AND DNN VIA PROXIMAL AVERAGE
4567	DETS: End-to-End Single-Stage Text-to-Speech via Hierarchical Diffusion Gan Models
3376	DF-VTON: Dense Flow Guided Virtual Try-On Network
6426	DGLP: INCORPORATING ORIENTATION INFORMATION FOR ENHANCED LINK PREDICTION IN DIRECTED GRAPHS
4940	DG-RAINDIFF: DEPTH-GUIDED DYNAMIC MESSAGE PASSING DIFFUSION MODEL FOR MIXTURE OF RAIN REMOVAL
5223	DIACORRECT: ERROR CORRECTION BACK-END FOR SPEAKER DIARIZATION
3819	DIAGNOSIS OF AUTISM SPECTRUM DISORDER BASED ON CONTRASTIVE FUNCTIONAL CONNECTIVITY GRAPH

<b>Paper Number</b>	<b>Paper Title</b>
	LEARNING NETWORK
1610	DIAGONALIZE INTEGRAL GRAPH BY DCT
8896	DIALCLIP: EMPOWERING CLIP AS MULTI-MODAL DIALOG RETRIEVER
7936	DIALOG MODELING IN AUDIOBOOK SYNTHESIS
4419	DIARIST: STREAMING SPEECH TRANSLATION WITH SPEAKER DIARIZATION
8291	DIB-X: FORMULATING EXPLAINABILITY PRINCIPLES FOR A SELF-EXPLAINABLE MODEL THROUGH INFORMATION THEORETIC LEARNING
4078	DICETRACK: LIGHTWEIGHT DICE CLASSIFICATION ON RESOURCE-CONSTRAINED PLATFORMS WITH OPTIMIZED DEEP LEARNING MODELS
5738	DIFFDUB: PERSON-GENERIC VISUAL DUBBING USING INPAINTING RENDERER WITH DIFFUSION AUTO-ENCODER
9504	DIFFERENTIABLE QUANTUM ARCHITECTURE SEARCH FOR JOB SHOP SCHEDULING PROBLEM
3427	Differentiable Resolution Compression and Alignment for Efficient Video Classification and Retrieval
11561	Differentiable Uncalibrated Imaging
4591	DIFFERENTIAL BEAMFORMING WITH NULL CONSTRAINTS FOR SPHERICAL MICROPHONE ARRAYS
8916	DIFFERENTIALLY PRIVATE FEDERATED FRANK-WOLFE
4622	DIFFEVENT: EVENT RESIDUAL DIFFUSION FOR IMAGE DEBLURRING
4946	DIFF-HOD: DIFFUSION MODEL FOR OBJECT DETECTION IN HAZY WEATHER CONDITIONS
8353	DiffRadar:High-quality mmWave Radar Perception with Diffusion Probabilistic Model
8995	DIFFRENT: A DIFFUSION MODEL FOR RECORDING ENVIRONMENT TRANSFER OF SPEECH
4374	DIFFSC: SEMANTIC COMMUNICATION FRAMEWORK WITH ENHANCED DENOISING THROUGH DIFFUSION PROBABILISTIC MODELS

<b>Paper Number</b>	<b>Paper Title</b>
8715	DIFFSTOCK: PROBABILISTIC RELATIONAL STOCK MARKET PREDICTIONS USING DIFFUSION MODELS
2673	DIFF-SV: A UNIFIED HIERARCHICAL FRAMEWORK FOR NOISE-ROBUST SPEAKER VERIFICATION USING SCORE-BASED DIFFUSION PROBABILISTIC MODELS
5717	DIFFUSION MODELS FOR AUDIO SEMANTIC COMMUNICATION
2729	DIFFUSION OPTIMISTIC LEARNING FOR MIN-MAX OPTIMIZATION
1622	DIFFUSION-BASED ADVERSARIAL PURIFICATION FOR ROBUST DEEP MRI RECONSTRUCTION
1572	DIFFUSION-BASED POSE REFINEMENT AND MULTI-HYPOTHESIS GENERATION FOR 3D HUMAN POSE ESTIMATION
3014	DIFFUSION-BASED SPEECH ENHANCEMENT IN MATCHED AND MISMATCHED CONDITIONS USING A HEUN-BASED SAMPLER
9097	Diffusion-based Speech Enhancement with a Weighted Generative-Supervised Learning Loss
3244	DIFFUSION-BASED SPEECH ENHANCEMENT WITH JOINT GENERATIVE AND PREDICTIVE DECODERS
1581	DiffusionInst: Diffusion Model for Instance Segmentation
8427	DIGITAL PATHOLOGY IMAGE DEBLURRING VIA LOCAL FOCUS QUALITY ASSESSMENT
8164	DIGITAL TASK-ORIENTED COMMUNICATION WITH HARDWARE-LIMITED TASK-BASED QUANTIZATION
3399	DI-MVS: LEARNING EFFICIENT MULTI-VIEW STEREO WITH DEPTH-AWARE ITERATIONS
2843	DIRECT POSITION DETERMINATION BY COVARIANCE-FITTING ON THE RIEMANNIAN MANIFOLD OF HERMITIAN POSITIVE DEFINITE MATRICES
7841	DIRECTED SCATTERING FOR KNOWLEDGE GRAPH-BASED CELLULAR SIGNALING ANALYSIS
3302	DIRECTIONAL GAIN BASED NOISE COVARIANCE MATRIX ESTIMATION FOR MVDR BEAMFORMING



<b>Paper Number</b>	<b>Paper Title</b>
8627	DISCOVERING MALICIOUS SIGNATURES IN SOFTWARE FROM STRUCTURAL INTERACTIONS
8077	DISCRETE AUDIO REPRESENTATION AS AN ALTERNATIVE TO MEL-SPECTROGRAMS FOR SPEAKER AND SPEECH RECOGNITION
4867	DISCRIMINANT PIXEL-DIFFERENCE VECTOR HASHING OF SPATIAL-TEMPORAL LOCAL BINARY PATTERNS FOR DYNAMIC TEXTURE RECOGNITION
10445	DISCRIMINATIVE FREQUENCY INFORMATION LEARNING FOR END-TO-END SPEECH ANTI-SPOOFING
7530	DISCRIMINATIVE SEMI-SUPERVISED FEATURE SELECTION VIA A CLASS-CREDIBLE PSEUDO-LABEL LEARNING FRAMEWORK
7483	DISCRIMINATIVE TRAINING OF VBX DIARIZATION
4985	DISENTANGLE ESTIMATION OF CAUSAL EFFECTS FROM CROSS-SILO DATA
5893	DISENTANGLED GRAPH REPRESENTATION WITH CONTRASTIVE LEARNING FOR RUMOR DETECTION
7159	DISENTANGLEMENT NETWORK: DISENTANGLE THE EMOTIONAL FEATURES FROM ACOUSTIC FEATURES FOR SPEECH EMOTION RECOGNITION
9762	DISENTANGLING THE SPECTRAL PROPERTIES OF THE HODGE LAPLACIAN: NOT ALL SMALL EIGENVALUES ARE EQUAL
3862	DISTILL VISION TRANSFORMERS TO CNNs VIA TEACHER COLLABORATION
1046	DISTILLING DISTRIBUTIONAL UNCERTAINTY FROM A GAUSSIAN PROCESS
7193	DISTILLING HUBERT WITH LSTMS VIA DECOUPLED KNOWLEDGE DISTILLATION
1589	Distributed Decision-Making for Community Structured Networks
11469	Distributed Self-Localization for Acoustic Transceiver Networks
11471	DISTRIBUTED SENSOR SELECTION FOR SPEECH ENHANCEMENT WITH ACOUSTIC SENSOR NETWORKS

<b>Paper Number</b>	<b>Paper Title</b>
8272	DISTRIBUTED STOCHASTIC CONTEXTUAL BANDITS FOR PROTEIN DRUG INTERACTION
6594	DISTRIBUTED VECTOR APPROXIMATE MESSAGE PASSING
6922	Distribution-aware Contrastive Learning for Robust Medical Image Segmentation
4445	DITW: a high-performance Deep-Independent Template-based Watermarking
8873	Diversifying Cross-Domain Few-shot Learning via Multimodal Image Editing
8744	Diversity based core-set selection for text-to-speech with linguistic and acoustic features
9751	Diversity-aware Buffer for Coping with Temporally Correlated Data Streams in Online Test-time Adaptation
1696	DJCM: A Deep Joint Cascade Model for Singing Voice Separation and Vocal Pitch Estimation
1140	DMEL: THE DIFFERENTIABLE LOG-MEL SPECTROGRAM AS A TRAINABLE LAYER IN NEURAL NETWORKS
4086	DMKD: IMPROVING FEATURE-BASED KNOWLEDGE DISTILLATION FOR OBJECT DETECTION VIA DUAL MASKING AUGMENTATION
1677	DMT: Comprehensive Distillation with Multiple Self-supervised Teachers
9226	DO LEARNED SPEECH SYMBOLS FOLLOW ZIPF'S LAW?
8951	DO SELF-SUPERVISED SPEECH AND LANGUAGE MODELS EXTRACT SIMILAR REPRESENTATIONS AS HUMAN BRAIN?
2469	DOA ESTIMATION FOR SWITCH-ELEMENT ARRAYS BASED ON SPARSE REPRESENTATION
8944	Does Audio Deepfake Detection Rely on Artifacts?
9584	DOES VIDEO SUMMARIZATION REQUIRE VIDEOS? QUANTIFYING THE EFFECTIVENESS OF LANGUAGE IN VIDEO SUMMARIZATION
9534	DOMAIN ADAPTIVE GRAPH CLASSIFICATION
7947	DOMAIN GENERALIZATION WITH FOURIER TRANSFORM AND SOFT THRESHOLDING

<b>Paper Number</b>	<b>Paper Title</b>
2180	Domain-adaptive and Subgroup-specific Cascaded Temperature Regression for Out-of-distribution Calibration
7639	DOMAIN-ADAPTIVE SEMANTIC SEGMENTATION EMERGES FROM VISION-LANGUAGE SUPERVISED DOMAIN-DEBIASED SELF-TRAINING
3013	DOMAINDIFF: BOOST OUT-OF-DISTRIBUTION GENERALIZATION WITH SYNTHETIC DATA
9181	DOMAIN-SLOT AWARE CONTRASTIVE LEARNING FOR IMPROVED DIALOGUE STATE TRACKING
3361	DOMAIN-WISE INVARIANT LEARNING FOR PANOPTIC SCENE GRAPH GENERATION
9885	DONE: DYNAMIC NEURAL REPRESENTATION VIA HYPERPLANE NEURAL ODE
10249	DOUBLE REVERSE REGULARIZATION NETWORK BASED ON SELF-KNOWLEDGE DISTILLATION FOR SAR OBJECT CLASSIFICATION
10106	DP-MAE: A DUAL-PATH MASKED AUTOENCODER BASED SELF-SUPERVISED LEARNING METHOD FOR ANOMALOUS SOUND DETECTION
7834	DPM-TSE: A Diffusion Probabilistic Model for Target Sound Extraction
8434	DRIVER SCANPATH PREDICTION BASED ON INVERSE REINFORCEMENT LEARNING
3403	Drop Sparse Convolution for 3D Object Detection
8927	DROPFL: CLIENT DROPOUT ATTACKS AGAINST FEDERATED LEARNING UNDER COMMUNICATION CONSTRAINTS
1405	DROPOUT MULTI-HEAD ATTENTION FOR SINGLE IMAGE SUPER-RESOLUTION
6781	DRSM: EFFICIENT NEURAL 4D DECOMPOSITION FOR DYNAMIC RECONSTRUCTION IN STATIONARY MONOCULAR CAMERAS
6411	DSIS: a novel (k,n) threshold deniable secret image sharing scheme with lossless recovery
7877	DT-NERF: DECOMPOSED TRIPLANE-HASH NEURAL RADIANCE FIELDS FOR HIGH-FIDELITY TALKING PORTRAIT

<b>Paper Number</b>	<b>Paper Title</b>
	SYNTHESIS
4733	DUAL CONTRASTIVE LEARNING GUIDED PATHOLOGICAL IMAGE RE-STAINING
1715	Dual Directional Complementary Gradient Fusion and Deep Refinement for Hyperspectral Image Super Resolution
8556	DUAL LEVEL INTENT-SLOT INTERACTION FOR IMPROVED MULTI-INTENT SPOKEN LANGUAGE UNDERSTANDING
4212	DUAL PARAMETER-EFFICIENT FINE-TUNING FOR SPEAKER REPRESENTATION VIA SPEAKER PROMPT TUNING AND ADAPTERS
7228	Dual Rank-1 Tensor Attention Module for Convolutional Neural Networks
7572	DUAL-CHANNEL UNLIMITED SAMPLING FOR BANDPASS SIGNALS
4810	DUAL-COLOR GRANULARITY ALIGNMENT FOR TEXT-BASED PERSON SEARCH
11867	DUAL-DOMAIN NEURAL NETWORKS FOR CLINICAL AND LOW-DOSE CBCT RECONSTRUCTION
6917	DualGCN-MIL: Whole Slide Image Classification Based on Double Relationship Graph Learning
6027	DUAL-MIX FOR CROSS-MODAL RETRIEVAL WITH NOISY LABELS
8112	DUAL-PATH MINIMUM-PHASE AND ALL-PASS DECOMPOSITION NETWORK FOR SINGLE CHANNEL SPEECH DEREVERBERATION
10264	DUAL-STREAM CONTRASTIVE PREDICTIVE NETWORK WITH JOINT HANDCRAFTED FEATURE VIEW FOR SAR SHIP CLASSIFICATION
4981	DUALVC 2: DYNAMIC MASKED CONVOLUTION FOR UNIFIED STREAMING AND NON-STREAMING VOICE CONVERSION
1354	DUNET: A ROBUST END-TO-END DEEP NEURAL NETWORK FRAMEWORK FOR IMBALANCED CLASSIFICATION
5628	DurIAN-E 2: Duration Informed Attention Network with Adaptive Variational Autoencoder and Adversarial Learning for Expressive Text-to-Speech Synthesis

<b>Paper Number</b>	<b>Paper Title</b>
1866	DURRNet: Deep Unfolded Single Image Reflection Removal Network with Joint Prior
9886	DUST: DUAL-GRAINED SYNTAX-AWARE TRANSFORMER NETWORK FOR CHINESE NAMED ENTITY RECOGNITION
8312	Dynamic ASR pathways: An Adaptive Masking Approach Towards Efficient Pruning of a Multilingual ASR Model
5556	DYNAMIC BANDWIDTH VARIATIONAL MODE DECOMPOSITION
5597	Dynamic Clustering and Cluster Contrastive Learning for Unsupervised Person Re-ID with Feature Distribution Alignment
5805	DYNAMIC DATA SAMPLER FOR CROSS-LANGUAGE TRANSFER LEARNING IN LARGE LANGUAGE MODELS
1897	DYNAMIC FREQUENCY DOMAIN GRAPH CONVOLUTIONAL NETWORK FOR TRAFFIC FORECASTING
1528	DYNAMIC LABEL SMOOTHING STRATEGY FOR BIOSIGNAL CLASSIFICATION
2030	DYNAMIC MODEL STRUCTURE ADJUSTMENT TO REALIZE QUANTUM CONTINUAL LEARNING BASED ON QUANTUM DATA
5536	DYNAMIC MULTI-SCALE CONTEXT AGGREGATION FOR CONVERSATIONAL ASPECT-BASED SENTIMENT QUADRUPLE ANALYSIS
2792	Dynamic Mutual-Activated Transformer for Human Motion Prediction
3575	DYNAMIC PRIVACY ALLOCATION FOR LOCALLY DIFFERENTIALLY PRIVATE FEDERATED LEARNING WITH COMPOSITE OBJECTIVES
7837	Dynamic random feature Gaussian Processes for Bayesian optimization of time-varying functions
3843	DYNAMIC REPLAY TRAINING FOR CLASS-INCREMENTAL LEARNING
7883	Dynamic Speech Emotion Recognition using a Conditional Neural Process
1259	DYNAMIC VIDEO FRAME INTERPOLATION WITH INTEGRATED DIFFICULTY PRE-ASSESSMENT

<b>Paper Number</b>	<b>Paper Title</b>
8147	DYNAMIC-SUPERB: TOWARDS A DYNAMIC, COLLABORATIVE, AND COMPREHENSIVE INSTRUCTION-TUNING BENCHMARK FOR SPEECH
2348	EARLY DIAGNOSING PARKINSON'S DISEASE VIA A DEEP LEARNING MODEL BASED ON AUGMENTED FACIAL EXPRESSION DATA
8902	ECHOCARDIOGRAPHY VIDEO SYNTHESIS FROM END DIASTOLIC SEMANTIC MAP VIA DIFFUSION MODEL
4936	ECIL-MU: EMBEDDING BASED CLASS INCREMENTAL LEARNING AND MACHINE UNLEARNING
3653	ECM-OPCC: EFFICIENT CONTEXT MODEL FOR OCTREE-BASED POINT CLOUD COMPRESSION
3069	EC-NAS: ENERGY CONSUMPTION AWARE TABULAR BENCHMARKS FOR NEURAL ARCHITECTURE SEARCH
8877	ECPNET: AN ENHANCED CURVE PERCEPTION NETWORK FOR LANE DETECTION
1864	Edge Attention Learning for Efficient Camouflaged Object Detection
4849	EDGE DEPLOYABLE DISTRIBUTED EVOLUTIONARY OPTIMIZATION BASED CALIBRATION METHOD FOR NEURAL QUANTIZATION
9959	EDM: Synthetic data from exemplar diffusion model improves non-communicable diseases detection
8157	ED-TTS: MULTI-SCALE EMOTION MODELING USING CROSS-DOMAIN EMOTION DIARIZATION FOR EMOTIONAL SPEECH SYNTHESIS
5644	EEG EMOTION RECOGNITION BASED ON DYNAMICAL GRAPH ATTENTION NETWORK
3659	EEG-BASED FAST AUDITORY ATTENTION DETECTION IN REAL-LIFE SCENARIOS USING TIME-FREQUENCY ATTENTION MECHANISM
8887	EFFECT OF BEAMPATTERN ON MATRIX COMPLETION WITH SPARSE ARRAYS
7386	EFFECT OF TARGET SIGNALS AND DELAYS ON SPATIALLY SELECTIVE ACTIVE NOISE CONTROL FOR OPEN-FITTING HEARABLES

<b>Paper Number</b>	<b>Paper Title</b>
3327	Effective Connectivity-based Multi-View Feature Learning Method for Dementia Diagnosis with fNIRS Signal
9338	EFFECTIVE IMAGE TAMPERING LOCALIZATION VIA ENHANCED TRANSFORMER AND CO-ATTENTION FUSION
9722	EFFECTIVE INTERNAL LANGUAGE MODEL TRAINING AND FUSION FOR FACTORIZED TRANSDUCER MODEL
10153	EFFICIENT 3D POSITION ESTIMATION IN BADMINTON SCENE
4394	Efficient Adapter Finetuning for Tail Languages in Streaming Multilingual ASR
8127	EFFICIENT ADAPTER TUNING OF PRE-TRAINED SPEECH MODELS FOR AUTOMATIC SPEAKER VERIFICATION
3941	Efficient Architecture Search for Real-time Instance Segmentation
9927	EFFICIENT BLACK-BOX SPEAKER VERIFICATION MODEL ADAPTATION WITH REPROGRAMMING AND BACKEND LEARNING
11547	EFFICIENT CODED MULTI-PARTY COMPUTATION AT EDGE NETWORKS
9734	EFFICIENT CONTENT RECONSTRUCTION FOR HIGH DYNAMIC RANGE IMAGING
4888	EFFICIENT FEDERATED LEARNING WITH SMOOTH AGGREGATION FOR NON-IID DATA FROM MULTIPLE EDGES
5937	EFFICIENT FUNCTIONAL LINK ADAPTIVE FILTERS BASED ON NEAREST KRONECKER PRODUCT DECOMPOSITION
1344	EFFICIENT FUSION OF DEPTH INFORMATION FOR DEFOCUS DEBLURRING
6814	EFFICIENT HIERARCHICAL STRIPE ATTENTION FOR LIGHTWEIGHT IMAGE SUPER-RESOLUTION
9176	EFFICIENT HIGH-PERFORMANCE BARK-SCALE NEURAL NETWORK FOR RESIDUAL ECHO AND NOISE SUPPRESSION
6098	EFFICIENT JOINT RECTIFICATION OF PHOTOMETRIC AND GEOMETRIC DISTORTIONS IN DOCUMENT IMAGES
8436	Efficient Learned Image Compression with Selective Kernel Residual Module and Channel-wise Causal Context Model

<b>Paper Number</b>	<b>Paper Title</b>
4304	EFFICIENT LEARNING ON SUCCESSIVE TEST TIME AUGMENTATION
5366	Efficient Multi-Channel Speech Enhancement with Spherical Harmonics Injection for Directional Encoding
8398	Efficient Personal Voice Activity Detection With Wake Word Reference Speech
9893	EFFICIENT POINT CLOUD ATTRIBUTE COMPRESSION FRAMEWORK USING ATTRIBUTE-GUIDED GRAPH FOURIER TRANSFORM
10159	EFFICIENT POINT CLOUD ATTRIBUTE COMPRESSION USING RICH PARALLELIZABLE CONTEXT MODEL
4236	Efficient Polyp Segmentation Via Integrity Learning
1448	Efficient PoseNet with Coarse to Fine Transformer
6898	EFFICIENT QUANTUM RECURRENT REINFORCEMENT LEARNING VIA QUANTUM RESERVOIR COMPUTING
3344	EFFICIENT SCENE TEXT IMAGE SUPER-RESOLUTION WITH SEMANTIC GUIDANCE
9000	EFFICIENT VIDEO AND AUDIO PROCESSING WITH LOIHI 2
6373	EiffHDR : AN EFFICIENT NETWORK FOR MULTI-EXPOSURE HIGH DYNAMIC RANGE IMAGING
6106	EIGENDECOMPOSITION-BASED SPATIAL-TEMPORAL ATTENTION FOR BRAIN COGNITIVE STATES IDENTIFICATION
5359	EK-NET:REAL-TIME SCENE TEXT DETECTION WITH EXPAND KERNEL DISTANCE
3974	ELECTROENCEPHALOGRAM HELPS FEW-SHOT LEARNING
8423	ELECTROENCEPHALOGRAM SENSOR DATA COMPRESSION USING AN ASYMMETRICAL SPARSE AUTOENCODER WITH A DISCRETE COSINE TRANSFORM LAYER
4670	ELECTROLARYNGEAL SPEECH INTELLIGIBILITY ENHANCEMENT THROUGH ROBUST LINGUISTIC ENCODERS
4190	Elevating Skeleton-Based Action Recognition with Efficient Multi-Modality Self-Supervision



<b>Paper Number</b>	<b>Paper Title</b>
4038	ELEVATING VISUAL PROMPTING IN TRANSFER LEARNING VIA PRUNED MODEL ENSEMBLES: NO RETRAIN, NO PAIN
2581	ELLIPSE DETECTION BASED ON CONTRAST-GUIDED ARC ENHANCEMENT
7399	ELLIPSE DETECTION BASED ON STRUCTURE-PRESERVING ANISOTROPIC EDGE EXTRACTION
7198	EMALG: AN ENHANCED MANDARIN LOMBARD GRID CORPUS WITH MEANINGFUL SENTENCES
1241	Embedded Feature Similarity Optimization with Specific Parameter Initialization for 2D/3D Medical Image Registration
8603	EMBEDDED GRAPH REPRESENTATION FOR INTER-FRAME CODING OF DYNAMIC MESHES
7109	EMOCONV-DIFF: DIFFUSION-BASED SPEECH EMOTION CONVERSION FOR NON-PARALLEL AND IN-THE-WILD DATA
4477	EMOHRNET: HIGH-RESOLUTION NEURAL NETWORK BASED SPEECH EMOTION RECOGNITION
4195	EMORED: A DATASET FOR RELATION EXTRACTION IN TEXTS WITH EMOTICONS
8212	EMOTALKER: EMOTIONALLY EDITABLE TALKING FACE GENERATION VIA DIFFUSION MODEL
1880	EMOTION NEURAL TRANSDUCER FOR FINE-GRAINED SPEECH EMOTION RECOGNITION
5697	EMOTION-ALIGNED CONTRASTIVE LEARNING BETWEEN IMAGES AND MUSIC
7429	EMOTION-AWARE CONTRASTIVE ADAPTATION NETWORK FOR SOURCE-FREE CROSS-CORPUS SPEECH EMOTION RECOGNITION
7301	EMOTVR: A HYBRID MODEL TO ESTIMATE CONTINUOUS-TIME AND CONTINUOUS-LEVEL EMOTION FROM ELECTROENCEPHALOGRAPHY
2624	Emphasized Non-Target Speaker Knowledge in Knowledge Distillation for Automatic Speaker Verification
4151	EMPLOYING REAL TRAINING DATA FOR DEEP NOISE SUPPRESSION

<b>Paper Number</b>	<b>Paper Title</b>
1706	EMPOWERING VISION-LANGUAGE MODELS FOR REASONING ABILITY THROUGH LARGE LANGUAGE MODELS
7944	ENABLING DEVICE CONTROL PLANNING CAPABILITIES OF SMALL LANGUAGE MODEL
8288	ENABLING ORIENTATION-FREE MMWAVE-BASED VITAL SIGN SENSING WITH MULTI-DOMAIN SIGNAL ANALYSIS
8643	ENABLING SECURE WIRELESS COMMUNICATIONS VIA MOVABLE ANTENNAS
7066	ENCLAP: COMBINING NEURAL AUDIO CODEC AND AUDIO-TEXT JOINT EMBEDDING FOR AUTOMATED AUDIO CAPTIONING
3713	Encoder-minimal and Decoder-minimal Framework for Remote Sensing Image Dehazing
4378	Encoding Seasonal Climate Predictions with Modular Neural Network
4338	ENCODING TIME AND ENERGY MODEL FOR SVT-AV1 BASED ON VIDEO COMPLEXITY
7447	END-TO-END LEARNING OF GAUSSIAN MIXTURE PROPOSALS USING DIFFERENTIABLE PARTICLE FILTERS AND NEURAL NETWORKS
11541	End-to-End Open Vocabulary Keyword Search With Multilingual Neural Representations
7924	END-TO-END PERSONALIZED CUFF-LESS BLOOD PRESSURE MONITORING USING ECG AND PPG SIGNALS
7223	End-to-end real time tracking of children's reading with pointer network
4203	end-to-end spatially-constrained multi-perspective fine-grained image captioning
8872	END-TO-END SPEECH RECOGNITION CONTEXTUALIZATION WITH LARGE LANGUAGE MODELS
5872	END-TO-END SPEECH TRANSLATION WITH MUTUAL KNOWLEDGE DISTILLATION
10323	ENERGY EFFICIENT WAKE-UP SOLUTION FOR LARGE-SCALE INTERNET OF UNDERWATER THINGS NETWORKS

<b>Paper Number</b>	<b>Paper Title</b>
7008	ENERGY-AWARE RESOLUTION SELECTION FOR PER-TITLE ENCODING
9680	ENERGY-BASED MODELS FOR SPEECH SYNTHESIS
4030	ENERGY-EFFICIENT DECENTRALIZED LEARNING VIA GRAPH SPARSIFICATION
5253	ENERGY-SAVING CELL-FREE MASSIVE MIMO PRECODERS WITH A PER-AP WIDEBAND KRONECKER CHANNEL MODEL
8176	ENGINEERING THE NEURAL COLLAPSE GEOMETRY OF SUPERVISED-CONTRASTIVE LOSS
6541	ENHANCED AXLE-BASED VEHICLE CLASSIFICATION USING ANGLE-BASED MICRO-DOPPLER SIGNATURE
4757	Enhanced Channel Estimation in mm-Wave MIMO Systems Leveraging Integrated Communication and Sensing
5998	ENHANCED COLOR PALETTE MODELING FOR LOSSLESS SCREEN CONTENT COMPRESSION
5000	ENHANCED DEEP REINFORCEMENT LEARNING FOR PARCEL SINGULATION IN NON-STATIONARY ENVIRONMENTS
8218	ENHANCED KPI ANOMALY DETECTION: AN UNSUPERVISED HYBRID MODEL WITH DYNAMIC THRESHOLD
3135	Enhanced low-rank and sparse Tucker decomposition for image completion
2486	ENHANCED SCREEN SHOOTING RESILIENT DOCUMENT WATERMARKING
3878	ENHANCED TRANSFER LEARNING WITH EFFICIENT MODELING AND ADAPTIVE FUSION OF KNOWLEDGE VIA PROMPT TUNING
1818	Enhanced Unsupervised Domain Adaptation with Dual-attention between Classification and Domain Alignment
4554	Enhancing Adversarial Robustness of DNNs via Weight Decorrelation in Training
4537	ENHANCING ADVERSARIAL TRAINING WITH PRIOR KNOWLEDGE DISTILLATION FOR ROBUST IMAGE COMPRESSION

<b>Paper Number</b>	<b>Paper Title</b>
2715	ENHANCING ADVERSARIAL TRANSFERABILITY IN OBJECT DETECTION WITH BIDIRECTIONAL FEATURE DISTORTION
9877	ENHANCING AOA ESTIMATION VIA PHASE MODELING OF BLUETOOTH 5 CTE SIGNALS
10039	ENHANCING ARGUMENTATIVE RELATION CLASSIFICATION BY MULTI-GRANULARITY RETRIEVAL AND HETEROGENEOUS GRAPH REASONING
5806	ENHANCING AUDIO GENERATION DIVERSITY WITH VISUAL INFORMATION
3288	ENHANCING AUDIO-VISUAL QUESTION ANSWERING WITH MISSING MODALITY VIA TRANS-MODAL ASSOCIATIVE LEARNING
4506	Enhancing Code-switching Speech Recognition with Interactive Language Biases
4736	ENHANCING CONVERSATION SMOOTHNESS IN LANGUAGE LEARNING CHATBOTS: AN EVALUATION OF GPT4 FOR ASR ERROR CORRECTION
6886	ENHANCING CROSS-DOMAIN DETECTION: ADAPTIVE CLASS-AWARE CONTRASTIVE TRANSFORMER
9647	ENHANCING DOCUMENT-LEVEL EVENT EXTRACTION VIA STRUCTURE-AWARE HETEROGENEOUS GRAPH WITH MULTI-GRANULARITY SUBSENTENCES
7752	ENHANCING END-TO-END CONVERSATIONAL SPEECH TRANSLATION THROUGH TARGET LANGUAGE CONTEXT UTILIZATION
4540	Enhancing Event Sequence Modeling with Contrastive Relational Inference
6729	ENHANCING EXPRESSIVENESS IN DANCE GENERATION VIA INTEGRATING FREQUENCY AND MUSIC STYLE INFORMATION
8551	ENHANCING GAN PERFORMANCE THROUGH NEURAL ARCHITECTURE SEARCH AND TENSOR DECOMPOSITION
3644	ENHANCING GENDER PRIVACY WITH PHOTO-REALISTIC FUSION OF DISENTANGLED SPATIAL SEGMENTS
8914	ENHANCING GENERALIZATION IN MEDICAL VISUAL QUESTION ANSWERING TASKS VIA GRADIENT-GUIDED

<b>Paper Number</b>	<b>Paper Title</b>
	MODEL PERTURBATION
2078	ENHANCING GENERALIZATION OF INVISIBLE FACIAL PRIVACY CLOAK VIA GRADIENT ACCUMULATION
3294	ENHANCING GENERATIVE ASPECT-BASED SENTIMENT ANALYSIS WITH RELATION-LEVEL SUPERVISION AND PROMPT
9585	ENHANCING HEALTHCARE WITH EOG: A NOVEL APPROACH TO SLEEP STAGE CLASSIFICATION
9929	ENHANCING HYPERSPECTRAL ANOMALY DETECTION BY DIFFERENCE-OF-CONVEX SPARSE ANOMALY MODELING
7903	ENHANCING IMAGE-TEXT MATCHING WITH ADAPTIVE FEATURE AGGREGATION
6144	ENHANCING LOW-LATENCY SPEAKER DIARIZATION WITH SPATIAL DICTIONARY LEARNING
4628	ENHANCING MULTILINGUAL SPEECH RECOGNITION THROUGH LANGUAGE PROMPT TUNING AND FRAME-LEVEL LANGUAGE ADAPTER
8247	ENHANCING MULTILINGUAL TTS WITH VOICE CONVERSION BASED DATA AUGMENTATION AND POSTERIOR EMBEDDING
1996	Enhancing Multi-task Models for Recommendation with Tensor Trace Norm
4817	ENHANCING NOISY LABEL LEARNING VIA UNSUPERVISED CONTRASTIVE LOSS WITH LABEL CORRECTION BASED ON PRIOR KNOWLEDGE
2038	ENHANCING NOTE-LEVEL SINGING TRANSCRIPTION MODEL WITH UNLABELED AND WEAKLY LABELED DATA
1139	ENHANCING PERFORMANCE OF COARSENEDED GRAPHS WITH GRADIENT-MATCHING
8610	ENHANCING PRE-TRAINED ASR SYSTEM FINE-TUNING FOR DYSARTHIC SPEECH RECOGNITION USING ADVERSARIAL DATA AUGMENTATION
8915	Enhancing Quantised End-to-End ASR Models via Personalisation
9036	ENHANCING REALISM IN 3D FACIAL ANIMATION USING CONFORMER-BASED GENERATION AND AUTOMATED

<b>Paper Number</b>	<b>Paper Title</b>
	POST-PROCESSING
4970	ENHANCING REINFORCEMENT LEARNING VIA CAUSALLY CORRECT INPUT IDENTIFICATION AND TARGETED INTERVENTION
4005	ENHANCING SEMANTIC COMMUNICATION WITH DEEP GENERATIVE MODELS: AN OVERVIEW
2412	ENHANCING SHORT- AND LONG-TERM SEA SURFACE TEMPERATURE FORECASTING WITH A STATIC AND DYNAMIC LEARNABLE PERSONALIZED GRAPH CONVOLUTION NETWORK
8670	ENHANCING SPATIAL AUDIO GENERATION WITH SOURCE SEPARATION AND CHANNEL PANNING LOSS
4416	ENHANCING SPEAKER DIARIZATION WITH LARGE LANGUAGE MODELS: A CONTEXTUAL BEAM SEARCH APPROACH
10118	ENHANCING STEGANOGRAPHY OF GENERATIVE IMAGE BASED ON IMAGE RETOUCHING
2122	ENHANCING TARGETED TRANSFERABILITY VIA FEATURE SPACE FINE-TUNING
2528	ENHANCING THE DOMAIN ROBUSTNESS OF SELF-SUPERVISED PRE-TRAINING WITH SYNTHETIC IMAGES
5910	Enhancing Two-stage Finetuning for Speech Emotion Recognition Using Adapters
4848	ENHANCING VIOLIN FINGERING GENERATION THROUGH AUDIO-SYMBOLIC FUSION
5124	ENRICHING MUSIC DESCRIPTIONS WITH A FINETUNED-LLM AND METADATA FOR TEXT-TO-MUSIC RETRIEVAL
2298	Entwined Inversion: Tune-Free Inversion for Real Image Faithful Reconstruction and Editing
2601	Environmental sound synthesis from vocal imitations and sound event labels
8608	EOFD-NET: EDGE OPTIMIZATION AND FEATURE DENOISING FOR WEAKLY SUPERVISED DEEP NUCLEI SEGMENTATION WITH PIONT ANNOTATIONS
6074	EPA: NEURAL COLLAPSE INSPIRED ROBUST OUT-OF-DISTRIBUTION DETECTOR

<b>Paper Number</b>	<b>Paper Title</b>
3264	ESA: EXPERT-AND-SAMPLES-AWARE INCREMENTAL LEARNING UNDER LONGTAIL DISTRIBUTION
5097	ESIHGNN: EVENT-STATE INTERACTIONS INFUSED HETEROGENEOUS GRAPH NEURAL NETWORK FOR CONVERSATIONAL EMOTION RECOGNITION
5922	ESTGN: ENHANCED SELF-MINED TEXT GUIDED SUPER-RESOLUTION NETWORK FOR SUPERIOR IMAGE SUPER RESOLUTION
8363	ESTIMATING DIRECTED SPECTRAL INFORMATION FLOW BETWEEN MULTI-RESOLUTION TIME SERIES
1824	ESTIMATING EXERCISE-INDUCED FATIGUE FROM THERMAL FACIAL IMAGES
3541	ESTIMATING SYMPTOMS AND CLINICAL SIGNS INSTEAD OF DISORDERS: THE PATH TOWARD THE CLINICAL USE OF VOICE AND SPEECH BIOMARKERS IN PSYCHIATRY
6224	ESTIMATION OF IMPULSE RESPONSES FOR A MOVING SOURCE USING OPTIMAL TRANSPORT REGULARIZATION
8962	ESTIMATION OF SPECTRAL LINES USING EXPECTATION PROPAGATION
8178	ESVC: COMBINING ADAPTIVE STYLE FUSION AND MULTI-LEVEL FEATURE DISENTANGLEMENT FOR EXPRESSIVE SINGING VOICE CONVERSION
7616	ETP: Learning Transferable ECG Representations via ECG-Text Pre-training
10028	Evaluation of an Improved ultrasonic imaging Helmet for observing Articulatory data
9457	EVIDENCE-AWARE MULTIMODAL CHINESE SOCIAL MEDIA RUMOR DETECTION
4370	EVOLUTION BACKCASTING OF EDGE FLOWS FROM PARTIAL OBSERVATIONS USING SIMPLICIAL VECTOR AUTOREGRESSIVE MODELS
7992	Exact classification of NMR spectra from NMR signals
2991	Exploiting A Quantum Multiple Kernel Learning Approach for Low-Resource Spoken Command Recognition
8693	EXPLOITING AUDIO-VISUAL FEATURES WITH PRETRAINED AV-HUBERT FOR MULTI-MODAL DYSARTHIC SPEECH

<b>Paper Number</b>	<b>Paper Title</b>
	RECONSTRUCTION
9972	EXPLOITING MODALITY-SPECIFIC FEATURES FOR MULTI-MODAL MANIPULATION DETECTION AND GROUNDING
2459	EXPLOITING SPATIAL-TEMPORAL DATA FOR SLEEP STAGE CLASSIFICATION VIA HYPERGRAPH LEARNING
4432	EXPLORATION OF VISUAL PROMPT IN GROUNDED PRE-TRAINED OPEN-SET DETECTION
9977	EXPLORING ADAPTERS WITH CONFORMERS FOR CHILDREN'S AUTOMATIC SPEECH RECOGNITION
3620	EXPLORING CONSISTENT SPATIO-TEMPORAL DISTORTION AND STABLE 3-D DCT COEFFICIENTS FOR ROBUST BLIND VIDEO WATERMARKING
6901	EXPLORING LABEL HIERARCHY IN DIALOGUE INTENT CLASSIFICATION
8714	Exploring large scale pre-trained models for robust machine anomalous sound detection
1594	Exploring Latent Cross-Channel Embedding for Accurate 3D Human Pose Reconstruction in a Diffusion Framework
7899	EXPLORING META INFORMATION FOR AUDIO-BASED ZERO-SHOT BIRD CLASSIFICATION
8239	EXPLORING MULTI-MODAL CONTROL IN MUSIC-DRIVEN DANCE GENERATION
7362	Exploring Object-centered External Knowledge for Fine-grained Video Paragraph Captioning
9665	EXPLORING PHONETIC CONTEXT-AWARE LIP-SYNC FOR TALKING FACE GENERATION
8528	EXPLORING SELF-EXPLAINABLE STREET-LEVEL IP GEOLOCATION WITH GRAPH INFORMATION BOTTLENECK
8198	EXPLORING SELF-SUPERVISED CONTRASTIVE LEARNING OF SPATIAL SOUND EVENT REPRESENTATION
8057	Exploring Soft Prompt Initialization Strategy for Few-shot Continual Text Classification
6220	EXPLORING SPATIO-TEMPORAL DISCRIMINATIVE CUES FOR GROUP ACTIVITY RECOGNITION VIA CONTRASTIVE LEARNING



<b>Paper Number</b>	<b>Paper Title</b>
6823	EXPLORING SPEECH RECOGNITION, TRANSLATION, AND UNDERSTANDING WITH DISCRETE SPEECH UNITS: A COMPARATIVE STUDY
4124	EXPLORING TARGETED UNIVERSAL ADVERSARIAL ATTACK FOR DEEP HASHING
7330	EXPLORING THE UTILITY OF CLIP PRIORS FOR VISUAL RELATIONSHIP PREDICTION
2794	Exponentially Consistent Nonparametric Clustering of Data Streams with Composite Distributions
8804	EXPRESSION DOMAIN TRANSLATION NETWORK FOR CROSS-DOMAIN HEAD REENACTMENT
9678	EXPRESSIVE ACOUSTIC GUITAR SOUND SYNTHESIS WITH AN INSTRUMENT-SPECIFIC INPUT REPRESENTATION AND DIFFUSION OUTPAINTING
10439	Extended Depth-of-Field Lensless Imaging Using an Optimized Radial Mask
5889	EXTENDING IMPLICIT NEURAL REPRESENTATIONS FOR TEXT-TO-IMAGE GENERATION
5531	EXTENDING LARGE LANGUAGE MODELS FOR SPEECH AND AUDIO CAPTIONING
3495	EXTENDING MULTILINGUAL ASR TO NEW LANGUAGES USING SUPPLEMENTARY ENCODER AND DECODER COMPONENTS
6956	EXTENDING MULTILINGUAL SPEECH SYNTHESIS TO 100+ LANGUAGES WITHOUT TRANSCRIBED DATA
9141	EXTENDING WHISPER WITH PROMPT TUNING TO TARGET-SPEAKER ASR
7391	EXTENSION OF CLIFFORD DATA REGRESSION METHODS FOR QUANTUM ERROR MITIGATION
3640	External Division of Two Proximity Operators: An Application to Signal Recovery with Structured Sparsity
7401	EXTREME ENCODER OUTPUT FRAME RATE REDUCTION: IMPROVING COMPUTATIONAL LATENCIES OF LARGE END-TO-END MODELS
7584	EXTREMELY LIGHT-WEIGHT LEARNING BASED LDR TO PQ HDR CONVERSION USING BERNSTEIN CURVES

<b>Paper Number</b>	<b>Paper Title</b>
6983	Extrinsic versus APP information feedback in turbo VEP MU-MIMO receivers: optimization via deep unfolding.
7123	EYE MOTION MATTERS FOR 3D FACE RECONSTRUCTION
1097	F1-EV SCORE: MEASURING THE LIKELIHOOD OF ESTIMATING A GOOD DECISION THRESHOLD FOR SEMI-SUPERVISED ANOMALY DETECTION
5129	F2GNN: AN ADAPTIVE FILTER WITH FEATURE SEGMENTATION FOR GRAPH-BASED FRAUD DETECTION
10073	FACE RECOGNITION USING LENSLESS CAMERA
9906	FACE RECONSTRUCTION FROM PARTIALLY LEAKED FACIAL EMBEDDINGS
6986	Facial Aesthetic Enhancement Network for Asian Faces Based on Differential Facial Aesthetic Activations
4553	FACIAL MICRO-MOTION-AWARE MIXUP FOR MICRO-EXPRESSION RECOGNITION
9358	FACILITATING MESSAGE PASSING WITH POTENTIAL LINKS FOR KNOWLEDGE GRAPH COMPLETION
8334	FACT-AWARE SUMMARIZATION WITH CONTRASTIVE LEARNING FOR FEW-SHOT DIALOGUE STATE TRACKING
5186	FAIRNESS-AWARE JOB SCHEDULING FOR MULTI-JOB FEDERATED LEARNING
3138	FALL PREDICTION BY A SPATIO-TEMPORAL MULTI-CHANNEL CAUSAL MODEL FROM WEARABLE SENSORS DATA
2032	FAMIM: A Novel Frequency-Domain Augmentation Masked Image Model Framework for Domain Generalizable Face Anti-Spoofing
8717	FAST ALGORITHM DESIGN FOR THE CONSTANT-ENVELOPE PRECODING IN MASSIVE MIMO COMMUNICATIONS WITH INTERFERENCE EXPLOITATION
1244	FAST ALIGNMENT ALGORITHM FOR CRYO-EM PARTICLE IMAGES BASED ON HARMONIC ANALYSIS
10139	FAST AND ACCURATE ROOT CAUSE ANALYSIS BASED ON SIGNALLING MESSAGES FOR 5G NETWORKS
6285	FAST AND EFFICIENT SEQUENTIAL RADAR PARAMETER ESTIMATION IN MIMO-OTFS SYSTEMS

<b>Paper Number</b>	<b>Paper Title</b>
3076	FAST AND PHYSICALLY ENRICHED DEEP NETWORK FOR JOINT LOW-LIGHT ENHANCEMENT AND IMAGE DEBLURRING
7623	FAST APPROXIMATION OF THE GENERALIZED SLICED-WASSERSTEIN DISTANCE
9455	Fast Cross-modality Knowledge Transfer via a Contextual Autoencoder Transformation
10424	Fast Dynamics of Brain-wide Patterns on Neuronal Oscillations
8365	FAST GRAPH-BASED DENOISING FOR POINT CLOUD COLOR INFORMATION
10084	Fast Intra mode prediction algorithms for SCBs in VVC SCC
4681	FAST PERSONALIZED TEXT TO IMAGE SYNTHESIS WITH ATTENTION INJECTION
9063	FAST TEST ERROR RATES FOR GRADIENT-BASED ALGORITHMS ON SEPARABLE DATA
2780	FASTGAT: SIMPLE AND EFFICIENT GRAPH ATTENTION NEURAL NETWORK WITH GLOBAL-AWARE ADAPTIVE COMPUTATIONAL NODE ATTENTION
7422	FASTINJECT: INJECTING UNPAIRED TEXT DATA INTO CTC-BASED ASR TRAINING
2958	FASTMANDARIN: EFFICIENT LOCAL MODELING FOR NATURAL MANDARIN SPEECH SYNTHESIS
3070	FAVANO: FEDERATED AVERAGING WITH ASYNCHRONOUS NODES
9084	FCC-MF: DETECTING VIOLENCE IN AUDIO-VISUAL CONTEXT WITH FRAME-WISE CLUSTER CONTRAST AND MODALITY-STAGE FLOODING
6991	FDA-MIMO Radar Using Ambiguity Function for Target Two-Dimensional Localization
5638	FDC-NERF: LEARNING POSE-FREE NEURAL RADIANCE FIELDS WITH FLOW-DEPTH CONSISTENCY
2329	FDIG: A Fine-grained Data Integration approach for Group Recommendation
3223	FDNET: A NOVEL MULTIVARIATE TIME SERIES CLASSIFICATION MODEL THROUGH FUSING FEATURE AND

<b>Paper Number</b>	<b>Paper Title</b>
	DIFFERENCE
10261	FEARLESS STEPS APOLLO: TEAM COMMUNICATIONS BASED DEVELOPMENT FOR SCIENCE, TECHNOLOGY, EDUCATION, AND HISTORICAL PRESERVATION
3378	Feature Mixing-based Active Learning for Multi-label Text Classification
2370	FEATURE-CONSTRAINED AND ATTENTION-CONDITIONED DISTILLATION LEARNING FOR VISUAL ANOMALY DETECTION
2163	Feature-Distribution Perturbation and Calibration for Generalized ReID
4357	FEDAQT: ACCURATE QUANTIZED TRAINING WITH FEDERATED LEARNING
2827	FEDERATED CINN CLUSTERING FOR ACCURATE CLUSTERED FEDERATED LEARNING
2915	Federated Dataset Dictionary Learning for Multi-Source Domain Adaptation
8131	Federated Learning of Tensor Generalized Linear Models with Low Separation Rank
1592	Federated Learning on Distributed Graphs considering Multiple Heterogeneities
8062	FEDERATED LEARNING UNDER RESTRICTED USER AVAILABILITY
9580	Federated Learning via Consensus Mechanism on Heterogeneous Data: A New Perspective on Convergence
3096	Federated Learning with Instance-Dependent Noisy Label
3894	FEDERATED PAC-BAYESIAN LEARNING ON NON-IID DATA
8565	FEDERATED QUANTUM MACHINE LEARNING WITH DIFFERENTIAL PRIVACY
8066	FedKA: Federated Knowledge Augmentation for Multi-Center Medical Image Segmentation on Non-IID Data
7814	FedLion: Faster Adaptive Federated Optimization with Fewer Communication
3073	FedMM: Federated Multi-Modal Learning with Modality Heterogeneity in Computational Pathology

<b>Paper Number</b>	<b>Paper Title</b>
9494	FED-SDS: ADAPTIVE STRUCTURED DYNAMIC SPARSITY FOR FEDERATED LEARNING UNDER HETEROGENEOUS CLIENTS
2635	FedSODA: Federated Cross-assessment and Dynamic Aggregation for Histopathology Segmentation
9939	FEWER-TOKEN NEURAL SPEECH CODEC WITH TIME-INVARIANT CODES
9182	FEW-SHOT ANOMALOUS SOUND DETECTION BASED ON ANOMALY MAP ESTIMATION USING PSEUDO ABNORMAL DATA
1794	FFT-BASED SELECTION AND OPTIMIZATION OF STATISTICS FOR ROBUST RECOGNITION OF SEVERELY CORRUPTED IMAGES
7455	FIBA: FEDERATED INVISIBLE BACKDOOR ATTACK
10134	Filamentary Convolution for Spoken Language Identification: A Brain-Inspired Approach
6465	FILTER-ENHANCED HYPERGRAPH TRANSFORMER FOR MULTI-BEHAVIOR SEQUENTIAL RECOMMENDATION
3274	FINCGAN: A GAN FRAMEWORK OF IMBALANCED NODE CLASSIFICATION ON HETEROGENEOUS GRAPH NEURAL NETWORK
6432	Finding Representative Sampling Subsets on Graphs via Submodularity
9259	FINE-GRAINED DISCREPANCY CONTRASTIVE LEARNING FOR ROBUST FAKE NEWS DETECTION
4855	FINE-GRAINED DISENTANGLED REPRESENTATION LEARNING FOR MULTIMODAL EMOTION RECOGNITION
7786	FINE-GRAINED ENGINE FAULT SOUND EVENT DETECTION USING MULTIMODAL SIGNALS
4048	FINE-GRAINED FEATURES ALIGNMENT AND FUSION FOR TEXT-VIDEO CROSS-MODAL RETRIEVAL
8402	Fine-Granularity Face Sketch Synthesis
6068	FINE-TUNE THE PRETRAINED ATST MODEL FOR SOUND EVENT DETECTION
7657	FINE-TUNING SELF-SUPERVISED MODELS FOR LANGUAGE IDENTIFICATION USING ORTHONORMAL CONSTRAINT

<b>Paper Number</b>	<b>Paper Title</b>
4442	FIRNET: FUNDAMENTAL FREQUENCY CONTROLLABLE FAST NEURAL VOCODER WITH TRAINABLE FINITE IMPULSE RESPONSE FILTER
8141	FIRST-SHOT UNSUPERVISED ANOMALOUS SOUND DETECTION WITH UNKNOWN ANOMALIES ESTIMATED BY METADATA-ASSISTED AUDIO GENERATION
7900	Fixed Inter-Neuron Covariability Induces Adversarial Robustness
9068	FLARE-FREE VISION: EMPOWERING UFORMER WITH DEPTH INSIGHTS
3925	FLATTENING SINGULAR VALUES OF FACTORIZED CONVOLUTION FOR MEDICAL IMAGES
1347	FLEXIBLE KEYWORD SPOTTING BASED ON HOMOGENEOUS AUDIO-TEXT EMBEDDING
1408	Flipping Consistent and Counterfactual Attention Network for Facial Expression Recognition
7029	FLOW DYNAMICS CORRECTION FOR ACTION RECOGNITION
7157	FOCUS FUSION NETWORK FOR VISIBLE AND INFRARED IMAGE FUSION
7588	Folding Attention: Memory and Power Optimization for On-Device Transformer-based Streaming Speech Recognition
6920	FOLLOWING THE EMBEDDING: IDENTIFYING TRANSITION PHENOMENA IN WAV2VEC 2.0 REPRESENTATIONS OF SPEECH AUDIO
3599	FORECASTING TORSIONAL RESONANCE IN ELECTRIC VEHICLES BY LEARNING A QUANTILE REGRESSOR
2314	Forgetting Private Textual Sequences in Language Models via Leave-One-Out Ensemble
8090	FOUNDATION MODEL ASSISTED AUTOMATIC SPEECH EMOTION RECOGNITION: TRANSCRIBING, ANNOTATING, AND AUGMENTING
4032	FOURIER DOMAIN APPROACH FOR GALAXY SPECTRA DECONTAMINATION AND DECONVOLUTION

<b>Paper Number</b>	<b>Paper Title</b>
8445	Fovea Transformer: Efficient Long-Context Modeling with Structured Fine-to-Coarse Attention
10228	FPGNET: SINGLE IMAGE DERAINING WITH HIGH-FREQUENCY CHANNEL AND FREQUENCY DOMAIN PRIOR GUIDANCE
4513	FPN WITH GMM BASED FEATURE ENHANCEMENT STRATEGY FOR OBJECT DETECTION IN REMOTE SENSING IMAGES
11496	Fractional Fourier Transform in Time Series Prediction
7747	FRACTURE ASSEMBLY WITH SEGMENTATION AND ITERATIVE REGISTRATION
6826	FRAME-LEVEL EMOTIONAL STATE ALIGNMENT METHOD FOR SPEECH EMOTION RECOGNITION
4260	Frame-to-Utterance Convergence: A Spectra-Temporal Approach for Unified Spoofing Detection
3291	FRAME-WISE STREAMING END-TO-END SPEAKER DIARIZATION WITH NON-AUTOREGRESSIVE SELF-ATTENTION-BASED ATTRACTORS
2757	FREETALKER: CONTROLLABLE SPEECH AND TEXT-DRIVEN GESTURE GENERATION BASED ON DIFFUSION MODELS FOR ENHANCED SPEAKER NATURALNESS
3045	FREEZE THE BACKBONES: A PARAMETER-EFFICIENT CONTRASTIVE APPROACH TO ROBUST MEDICAL VISION-LANGUAGE PRE-TRAINING
4175	FREGRAD: LIGHTWEIGHT AND FAST FREQUENCY-AWARE DIFFUSION VOCODER
6963	FREMAX: A SIMPLE METHOD TOWARDS TRULY SECURE GENERATIVE LINGUISTIC STEGANOGRAPHY
7494	FREQ2TIME: WEAKLY SUPERVISED LEARNING OF CAMERA-BASED RPPG FROM HEART RATE
7202	FREQUENCY ANALYSIS AND FILTER DESIGN FOR DIRECTED GRAPHS WITH POLAR DECOMPOSITION
1736	FREQUENCY AWARE AND GRAPH FUSION NETWORK FOR POLYP SEGMENTATION
6946	FREQUENCY ESTIMATION VIA SUB-NYQUIST UNLIMITED SAMPLING

<b>Paper Number</b>	<b>Paper Title</b>
8934	FREQUENCY MASKING FOR UNIVERSAL DEEPFAKE DETECTION
8446	FREQUENCY-DOMAIN SIGNAL RECONSTRUCTION FOR DYNAMIC TIME-DOMAIN WEIGHTING HYBRID PRECODING WITH BEAM SQUINT
2470	Friends to Help: Saving Federated Learning from Client Dropout
9284	FROM COARSE TO FINE: EFFICIENT TRAINING FOR AUDIO SPECTROGRAM TRANSFORMERS
2725	FROM CONVOLUTIONAL SPARSE CODING TO *-NMF FACTORIZATION OF TIME-FREQUENCY COEFFICIENTS
7234	FROM GAME THEORY TO VISUAL RECOGNITION: ADVANCING DNN ROBUSTNESS
7963	FROM RIR TO BRIR: A SPARSE RECOVERY BEAMFORMING APPROACH FOR VIRTUAL BINAURAL SOUND RENDERING
3633	FSD: AN INITIAL CHINESE DATASET FOR FAKE SONG DETECTION
3820	FSPEN: AN ULTRA-LIGHTWEIGHT NETWORK FOR REAL TIME SPEECH ENAHNCMENT
3741	FUNCODEC: A FUNDAMENTAL, REPRODUCIBLE AND INTEGRABLE OPEN-SOURCE TOOLKIT FOR NEURAL SPEECH CODEC
4576	Functional Emotion Transformer for EEG-assisted Cross-Modal Emotion Recognition
7833	Functional Invariants to Watermark Large Transformers
8377	FUNCTIONALLY SIMILAR MULTI-LABEL KNOWLEDGE DISTILLATION
6308	FUNDAMENTAL LIMITS OF DIRECTION FINDING IN DISTRIBUTED ARRAYS EXPLOITING AUXILIARY SOURCES
9304	FUNDAMENTAL PERFORMANCE BOUNDS FOR CARRIER PHASE POSITIONING IN LEO-PNT SYSTEM
2621	FUR-API: DATASET AND BASELINES TOWARD REALISTIC API ANOMALY DETECTION
3867	FURTHER RESULTS ON THE DESIGN OF REAL-VALUED WIDEBAND BEAMFORMERS USING ADAPTIVE-ARRAY-THEORY-INSPIRED WEIGHTED LEAST SQUARES



<b>Paper Number</b>	<b>Paper Title</b>
9364	FUSDOM: COMBINING IN-DOMAIN AND OUT-OF-DOMAIN KNOWLEDGE FOR CONTINUOUS SELF-SUPERVISED LEARNING
2550	Fusing Modality-Specific Representations and Decisions for Multimodal Emotion Recognition
7241	FUSING MULTI-LEVEL FEATURES FROM AUDIO AND CONTEXTUAL SENTENCE EMBEDDING FROM TEXT FOR INTERVIEW-BASED DEPRESSION DETECTION
5573	FUSING STRUCTURE AND APPEARANCE FEATURES IN FACIAL EXPRESSION RECOGNITION TRANSFORMER
9878	FUSION OF AUDIO AND VISUAL EMBEDDINGS FOR SOUND EVENT LOCALIZATION AND DETECTION
7948	Fusion of Multi-resolution Seismic Tomography Maps with Physics-informed Probability Graphical Models
2150	FuzzLLM: A Novel and Universal Fuzzing Framework for Proactively Discovering Jailbreak Vulnerabilities in Large Language Models
4738	FW-SHAPLEY: REAL-TIME ESTIMATION OF WEIGHTED SHAPLEY VALUES
1615	G2G: GENERALIZED LEARNING BY CROSS-DOMAIN KNOWLEDGE TRANSFER FOR FEDERATED DOMAIN GENERALIZATION
1746	G2PU: Grapheme-to-Phoneme Transducer with Speech Units
7756	GAMAFLOW: ESTIMATING 3D SCENE FLOW VIA GROUPED ATTENTION AND GLOBAL MOTION AGGREGATION
3536	GaP-aug: Gamma Patch-Wise Correction Augmentation Method for Respiratory Sound Classification
3528	GASS: GENERALIZING AUDIO SOURCE SEPARATION WITH LARGE-SCALE DATA
8078	GBSD: GENERATIVE BOKEH WITH STAGE DIFFUSION
9846	GCC-PHAT RE-IMAGINED - A U-NET FILTER FOR AUDIO TDOA PEAK-SELECTION
6454	GCIA: A BLACK-BOX GRAPH INJECTION ATTACK METHOD VIA GRAPH CONTRASTIVE LEARNING

<b>Paper Number</b>	<b>Paper Title</b>
1608	GEemo-CLAP: Gender-Attribute-Enhanced Contrastive Language-Audio Pretraining for Accurate Speech Emotion Recognition
4147	GENEFORMER: LEARNED GENE COMPRESSION USING TRANSFORMER-BASED CONTEXT MODELING
5185	General Phrase Debiaser: Debiasing Masked Language Models at a Multi-Token Level
11876	GENERAL SPEECH RESTORATION USING TWO-STAGE GENERATIVE ADVERSARIAL NETWORKS
9367	GENERALIZABLE TWO-BRANCH FRAMEWORK FOR IMAGE CLASS-INCREMENTAL LEARNING
7863	Generalization of self-supervised learning-based representations for cross-domain speech emotion recognition
1564	GENERALIZED DETERMINISTIC-RANDOM TRADEOFF OF INTEGRATED SENSING AND COMMUNICATIONS: THE SENSING-OPTIMAL OPERATING POINT
8011	GENERALIZED HOLE-FILLING STRATEGY FOR OVERLAPPING HOLE-EXISTING COPRIME ARRAYS FOR DOA ESTIMATION
7848	Generalized Multi-Source Inference for Text Conditioned Music Diffusion Models
5472	GENERALIZED SPECAUGMENT VIA MULTI-RECTANGLE INVERSE MASKING FOR ACOUSTIC SCENE CLASSIFICATION
7168	Generalized Uncertainty-Based Evidential Fusion with Hybrid Multi-Head Attention for Weak-Supervised Temporal Action Localization
2759	Generating High-quality Adversarial Examples with Universal Perturbation-Based Adaptive Network and Improved Perceptual Loss
8932	GENERATING PERSONA-AWARE EMPATHETIC RESPONSES WITH RETRIEVAL-AUGMENTED PROMPT LEARNING
10023	GENERATING STEREOPHONIC MUSIC WITH SINGLE-STAGE LANGUAGE MODELS
7649	GENERATION OR REPLICATION: AUSCULTATING AUDIO LATENT DIFFUSION MODELS

<b>Paper Number</b>	<b>Paper Title</b>
9530	Generation-based Target Speech Extraction with Speech Discretization and Vocoder
2445	GENERATIVE AI-AIDED JOINT TRAINING-FREE SECURE SEMANTIC COMMUNICATIONS VIA MULTI-MODAL PROMPTS
5119	GENERATIVE CONTEXT-AWARE FINE-TUNING OF SELF-SUPERVISED SPEECH MODELS
8060	GENERATIVE DE-QUANTIZATION FOR NEURAL SPEECH CODEC VIA LATENT DIFFUSION
4244	Generative Extension Positive Pairs and Improving Sample Selection Based on Contrastive Learning for Unsupervised Person Re-identification
7557	GEODESIC INTERPOLATION OF FRAME-WISE SPEAKER EMBEDDINGS FOR THE DIARIZATION OF MEETING SCENARIOS
8937	Geometry Compression Artifact Removal for V-PCC over a Wide Bitrate Range
3457	GEOMETRY-CORRECTED GEODESIC MOTION MODELING WITH PER-FRAME CAMERA MOTION FOR 360-DEGREE VIDEO COMPRESSION
8582	GESTURE GENERATION VIA DIFFUSION MODEL WITH ATTENTION MECHANISM
11475	GFANC-Kalman: Generative Fixed-Filter Active Noise Control with CNN-Kalman Filtering
9257	GFMAE: Self-Supervised GNN-Free Masked AutoEncoders
5005	GI-PIP: DO WE REQUIRE IMPRACTICAL AUXILIARY DATASET FOR GRADIENT INVERSION ATTACKS?
7033	GLA-GRAD: A GRIFFIN-LIM EXTENDED WAVEFORM GENERATION DIFFUSION MODEL
3562	GLANCE, FOCUS AND REFINEMENT NETWORK FOR REMOTE SENSING CHANGE DETECTION
6216	GLANCING FUTURE FOR SIMULTANEOUS MACHINE TRANSLATION
1497	GLAND INSTANCE SEGMENTATION BY FULL RESOLUTION MULTI-SCALE DILATION RESIDUAL NETWORKS

<b>Paper Number</b>	<b>Paper Title</b>
10177	GLAND SEGMENTATION VIA DUAL ENCODERS AND BOUNDARY-ENHANCED ATTENTION
3535	GLMAE: GRAPH REPRESENTATION LEARNING METHOD COMBINING GENERATIVE LEARNING AND MASKING AUTOENCODER
5172	GLMB 3D SPEAKER TRACKING WITH VIDEO-ASSISTED MULTI-CHANNEL AUDIO OPTIMIZATION FUNCTIONS
2165	Global Convergence of Alternating Direction Method of Multipliers for Invex Objective Losses
1668	GLOBAL OPTIMIZATION OF ACTIVE RIS IN LINEAR TIME
11550	Global Optimization of Long-Term Average Proportional Fair Throughput via Convex Reformulation
3239	Globally Optimal Beamforming Design for Integrated Sensing and Communication Systems
1938	GLOCAL CASCADING NETWORK FOR TOPIC ENHANCED VISUAL STORYTELLING
8043	GMM-RESNET2: ENSEMBLE OF GROUP RESNET NETWORKS FOR SYNTHETIC SPEECH DETECTION
7184	GMM-RESNEXT: COMBINING GENERATIVE AND DISCRIMINATIVE MODELS FOR SPEAKER VERIFICATION
6173	GMTR: Graph Matching Transformers
6861	GM-VRC: SEMANTIC TOPOLOGICAL DATA ENSEMBLE APPROACH FOR EEG SIGNAL CLASSIFICATION
7878	GPT-4 DRIVEN CINEMATIC MUSIC GENERATION THROUGH TEXT PROCESSING
1691	GPTCN: GATED PARALLEL TRANSFORMER CONVOLUTIONAL NETWORKS FOR DOWNSTREAM-TASK USER REPRESENTATION LEARNING ON APP USAGE
4300	GR0: Self-supervised Global Representation Learning for Zero-shot Voice Conversion
4538	GRADIENT AND BRIGHTNESS GUIDED LOW-LIGHT ENHANCEMENT WITH ATTENTION-BASED SELF-PACED LEARNING
8532	Gradient Inversion Attacks on Acoustic Signals: Revealing Security Risks in Audio Recognition Systems

<b>Paper Number</b>	<b>Paper Title</b>
6947	Gradient Reactivation Enhanced Causal Attention for Out-Of-Distribution Generalizable Graph Classification
5876	GRADIENT WEIGHTING FOR SPEAKER VERIFICATION IN EXTREMELY LOW SIGNAL-TO-NOISE RATIO
3421	GRADIENT-AWARE LOGIT ADJUSTMENT LOSS FOR LONG-TAILED CLASSIFIER
6847	GRADIENT-BASED DIMENSIONALITY REDUCTION FOR SPEECH EMOTION RECOGNITION USING DEEP NETWORKS
7633	GRADUALLY SPATIO-TEMPORAL FEATURE ACTIVATION FOR TARGET TRACKING
3832	Granger Connectivity Analysis as a Block-Term Tensor Regression for eSport Players
11514	GRAPH ATTENTION FOR AUTOMATED AUDIO CAPTIONING
8029	Graph Convolutional Neural Networks in the Companion Model
8289	GRAPH IDENTIFICATION AND UPPER CONFIDENCE EVALUATION FOR CAUSAL BANDITS WITH LINEAR MODELS
6174	GRAPH LOCAL-SMOOTH DICTIONARY LEARNING
8660	Graph Networks Stand Strong: Enhancing Robustness via Stability Constraints
9469	GRAPH NEURAL NETWORKS ARE MORE POWERFUL THAN WE THINK
8509	Graph Signal Processing: The 2D Companion Model
1086	GRAPH-AWARE MULTI-VIEW FUSION FOR RUMOR DETECTION ON SOCIAL MEDIA
8882	Graph-based Environment Representation for Vision-and-Language Navigation in Continuous Environments
7849	GRAPH-BASED PERMUTATION PATTERNS FOR THE ANALYSIS OF TASK-RELATED FMRI SIGNALS ON DTI NETWORKS IN MILD COGNITIVE IMPAIRMENT
8152	Graph-enhanced Hybrid Sampling for Multi-armed Bandit Recommendation

<b>Paper Number</b>	<b>Paper Title</b>
5839	Graphical Inference in Non-Markovian Linear-Gaussian State-space Models
11455	GRAPHON POOLING FOR REDUCING DIMENSIONALITY OF SIGNALS AND CONVOLUTIONAL OPERATORS ON GRAPHS
3907	Gravitated Latent Space Loss Generated by Metric Tensor for High-Dynamic Range Imaging
9389	GRIDLESS PARAMETER ESTIMATION IN PARTLY CALIBRATED RECTANGULAR ARRAYS
6486	GROUNDED-INSTRUCT-PIX2PIX: IMPROVING INSTRUCTION BASED IMAGE EDITING WITH AUTOMATIC TARGET GROUNDING
11462	GROUSE: A TASK AND MODEL AGNOSTIC WAVELET-DRIVEN FRAMEWORK FOR MEDICAL IMAGING
3943	G-SharP: Globally Shared Kernel with Pruning for Efficient CNNs
8780	GSTNet: Gait Spatio-Temporal Network for Gait Recognition Using Millimeter-Wave Radar
6951	GTCRN: A Speech Enhancement Model Requiring Ultralow Computational Resources
1369	GuessKT: Improving Knowledge Tracing via Considering Guess Behaviors
1960	GUIDED CIRCULAR DECOMPOSITION AND CROSS-MODAL RECOMBINATION FOR MULTIMODAL SENTIMENT ANALYSIS
4826	HADGEO: IMAGE BASED 3-DOF CROSS-VIEW GEO-LOCALIZATION WITH HARD SAMPLE MINING
5575	HAFFORMER: A HIERARCHICAL ATTENTION-FREE FRAMEWORK FOR ALZHEIMER'S DISEASE DETECTION FROM SPONTANEOUS SPEECH
4198	HAFORMER: HETEROGENEOUS AGGREGATION TRANSFORMER FOR SINGLE IMAGE DERAINING
11545	HALF-INVERTED ARRAY DESIGN SCHEME FOR LARGE HOLE-FREE FOURTH-ORDER DIFFERENCE CO-ARRAYS
9523	HALTINGVT: ADAPTIVE TOKEN HALTING TRANSFORMER FOR EFFICIENT VIDEO RECOGNITION

<b>Paper Number</b>	<b>Paper Title</b>
7307	HARDWARE IMPAIRMENTS-AWARE DESIGN OF NONCOHERENT GRASSMANNIAN CONSTELLATIONS
7967	Hardware-Algorithm Co-design Enabling Processing-in-Pixel-in-Memory ( $P^2M$ ) for Neuromorphic Vision Sensors
7879	HARDWARE-LIMITED TIME CONSTANT ESTIMATION USING A WEIGHTED LINEAR REGRESSION
7887	Harmonic Retrieval for Non-Circular Coherent Signals via Double Decoupled Atomic Norm Minimization
6916	HARNESSING THE POWER OF LARGE VISION LANGUAGE MODELS FOR SYNTHETIC IMAGE DETECTION
7761	HAROOD: HUMAN ACTIVITY CLASSIFICATION AND OUT-OF-DISTRIBUTION DETECTION WITH SHORT-RANGE FMCW RADAR
8664	HAZY REMOTE SENSING IMAGES SEMANTIC SEGMENTATION FOR WEAKLY ANNOTATION BASED ON SALIENCY-AWARE ALIGNMENT STRATEGY
6116	HDPNeRF: Hybrid Depth Priors for Neural Radiance Fields from Sparse Input Views
1738	HDRTVFormer: Efficient SDRTV-to-HDRTV via Affine Transformation and Spatial-aware Transformer
10426	HEALTHY AGING IS MARKED BY ENTROPY REDUCTION IN CORTICAL SPONTANEOUS ACTIVITY
2516	HEARING LOSS DETECTION FROM FACIAL EXPRESSIONS IN ONE-ON-ONE CONVERSATIONS
7908	Heart Rate Variability Estimation with Dynamic Fine Filtering and Global-Local Context Outlier Removal
8494	HEAR-YOUR-ACTION: HUMAN ACTION RECOGNITION BY ULTRASOUND ACTIVE SENSING
5399	HENET: HYPERBOLIC-BASED ENCODER-DECODER NETWORK FOR WORD SPOTTING IN HISTORICAL MONGOLIAN DOCUMENTS
7107	HETEROGENEOUS FACE RECOGNITION USING DOMAIN INVARIANT UNITS
4251	HEURISTIC-DRIVEN, TYPE-SPECIFIC EMBEDDING IN PARALLEL SPACES FOR ENHANCING KNOWLEDGE GRAPH REASONING

<b>Paper Number</b>	<b>Paper Title</b>
9687	Hierarchical Attacks on Large-Scale Graph Neural Networks
4839	Hierarchical cross-modality knowledge transfer with sinkhorn attention for CTC-based ASR
3529	HIERARCHICAL EMOTION PREDICTION AND CONTROL IN TEXT-TO-SPEECH SYNTHESIS
8256	HIERARCHICAL HOME ACTION UNDERSTANDING WITH IMPLICIT AND EXPLICIT PRIOR KNOWLEDGE
9753	HIERARCHICAL METADATA INFORMATION CONSTRAINED SELF-SUPERVISED LEARNING FOR ANOMALOUS SOUND DETECTION UNDER DOMAIN SHIFT
2708	HIERARCHICAL SPEAKER REPRESENTATION FOR TARGET SPEAKER EXTRACTION
2730	HIERARCHICAL VAE BASED SEMANTIC COMMUNICATIONS FOR POMDP TASKS
8277	High Accuracy Device Localization in Indoor mmWave Networks Exploiting Channel Sparsity and Virtual Anchor Mapping
7349	HIGH RESOLUTION GUITAR TRANSCRIPTION VIA DOMAIN ADAPTATION
3027	HIGH RESOLUTION IMAGE QUALITY DATABASE
5114	HIGH-ACCURACY ANXIETY DISORDER IDENTIFICATION THROUGH SUBSPACE-ENHANCED HYPERGRAPH NEURAL NETWORK
8126	Higher Order Multiple Graph Filtering for Structured Graph Learning
11465	High-Fidelity and Pitch-Controllable Neural Vocoder Based on Unified Source-Filter Networks
4276	High-Fidelity Speech Synthesis with Minimal Supervision: All Using Diffusion Models
1380	Highlight removal network based on an improved dichromatic reflection model
7319	HIGH-ORDER TENSOR POOLING WITH ATTENTION FOR ACTION RECOGNITION
4701	HIGH-RESOLUTION THROUGH-WALL IMAGING USING DATA FUSION AND REASONING



<b>Paper Number</b>	<b>Paper Title</b>
3729	HIM: DISCOVERING IMPLICIT RELATIONSHIPS IN HETEROGENEOUS SOCIAL NETWORKS
2379	HINT-ENHANCED IN-CONTEXT LEARNING WAKES LARGE LANGUAGE MODELS UP FOR KNOWLEDGE-INTENSIVE TASKS
5535	HIQ: ONE-SHOT NETWORK QUANTIZATION FOR HISTOPATHOLOGICAL IMAGE CLASSIFICATION
11512	HISTORICAL AUDIO SEARCH AND PRESERVATION: FINDING WALDO WITHIN THE FEARLESS STEPS APOLLO 11 NATURALISTIC AUDIO CORPUS
8868	HLS-FGVC: HIERARCHICAL LABEL SEMANTICS ENHANCED FINE-GRAINED VISUAL CLASSIFICATION
3491	HM-CONFORMER: A CONFORMER-BASED AUDIO DEEPFAKE DETECTION SYSTEM WITH HIERARCHICAL POOLING AND MULTI-LEVEL CLASSIFICATION TOKEN AGGREGATION METHODS
8072	HMM-Based CSI Embedding For Trajectory Recovery from RSS Measurements of Non-Cooperative Devices
1450	HMNet: Hierarchical Microscale-aware Network for Infrared Small Target Detection
7827	HODGE-AWARE CONTRASTIVE LEARNING
8851	HOICS: Zero-Shot HOI Detection via Compatibility Self-Learning
4366	HOT-FIXING WAKE WORD RECOGNITION FOR END-TO-END ASR VIA NEURAL MODEL REPROGRAMMING
2684	HOURLASS-AVSR: DOWN-UP SAMPLING-BASED COMPUTATIONAL EFFICIENCY MODEL FOR AUDIO-VISUAL SPEECH RECOGNITION
1379	How Can Personalized Context Help? Exploring Joint Retrieval of Passage and Personalized Context
4802	HOW DOES END-TO-END SPEECH RECOGNITION TRAINING IMPACT SPEECH ENHANCEMENT ARTIFACTS?
7272	HOW SECURE IS THE TIME-MODULATED ARRAY-ENABLED OFDM DIRECTIONAL MODULATION?
10254	HOW TO BRIDGE GRAPH AND SEQUENCE PATTERNS IN SESSION-BASED RECOMMENDATION? A SELF-SUPERVISED

<b>Paper Number</b>	<b>Paper Title</b>
	METHOD
10437	HOW TO DISTURB NETWORK RECONNAISSANCE: A MOVING TARGET DEFENSE APPROACH BASED ON DEEP REINFORCEMENT LEARNING
7537	HRTF Recommendation Based on the Predicted Binaural Colouration Model
7269	HUBERTOPIC: ENHANCING SEMANTIC REPRESENTATION OF HUBERT THROUGH SELF-SUPERVISION UTILIZING TOPIC MODEL
1709	Human Guided Cross-Modal Reasoning with Semantic Attention Learning for Visual Question Answering
8143	HUMAN MOTION CAPTURE DATA SEGMENTATION BASED ON ST-GCN
4495	HUMAN MOTION GENERATION VIA CONDITIONED GMVAE WITH TUNET
7094	Human Perception-Guided Meta-training for Few-shot NeRF
2267	HUMTRANS: A NOVEL OPEN-SOURCE DATASET FOR HUMMING MELODY TRANSCRIPTION AND BEYOND
3001	Hybrid Attention Time-Frequency Analysis Network for Single-Channel Speech Enhancement
1874	HYBRID CONVOLUTION-TRANSFORMER FOR LIGHTWEIGHT SINGLE IMAGE SUPER-RESOLUTION
2264	HYBRID DOMAIN LEARNING TOWARDS LIGHT FIELD SPATIAL SUPER-RESOLUTION USING HETEROGENEOUS IMAGING
5132	Hybrid Module with Multiple Receptive Fields and Self-Attention Layers for Medical Image Segmentation
5107	HYPERBOLIC DIFFUSION PROCRUSTES ANALYSIS FOR INTRINSIC REPRESENTATION OF HIERARCHICAL DATA SETS
7816	HYPERBOLIC DISTANCE-BASED SPEECH SEPARATION
6786	HyperGANStrument: Instrument Sound Synthesis and Editing with Pitch-Invariant Hypernetworks
9342	HYPERGRAPH TRANSFORMER FOR SEMI-SUPERVISED CLASSIFICATION

<b>Paper Number</b>	<b>Paper Title</b>
2732	HYPERGRAPH-ENHANCED SELF-SUPERVISED ROBUST GRAPH LEARNING FOR SOCIAL RECOMMENDATION
8986	Hypergraph-MLP: Learning On Hypergraphs Without Message Passing
11518	HYPERPIXELS: FLEXIBLE 4D OVER-SEGMENTATION FOR DENSE AND SPARSE LIGHT FIELDS
6267	Hyperspectral Image Reconstruction using Hierarchical Neural Architecture Search from a Snapshot Image
11927	HYPERSPSCTRAL RECONSTRUCTION OF SKIN THROUGH FUSION OF SCATTERING TRANSFORM FEATURES
11895	HYPERSPSCTRAL SKIN VISION CHALLENGE: CAN YOUR CAMERA SEE BEYOND YOUR SKIN?
11914	HYSAT++: HYBRID SPECTRAL-WISE ATTENTION TRANSFORMER FOR SKIN SPECTRAL RECONSTRUCTION
9475	HYSENSE: HYBRID EVENT OCCURRENCE DETECTION METHOD FOR IOT DEVICES
5699	HYSTOC: OBTAINING WORD CONFIDENCES FOR FUSION OF END-TO-END ASR SYSTEMS
2442	I3FDM: Iris Inpainting via Inverse Fusion of Diffusion Models
11961	ICASSP 2024 Auditory EEG Decoding Challenge
11865	ICASSP 2024 SPEECH SIGNAL IMPROVEMENT CHALLENGE
11900	ICMC-ASR: THE ICASSP 2024 IN-CAR MULTI-CHANNEL AUTOMATIC SPEECH RECOGNITION CHALLENGE
2308	IDENTIFIABILITY ANALYSIS OF SENSOR ARRAYS WITH SENSORS OFF HALF-WAVELENGTH GRID
10012	IDENTIFIABILITY STUDY OF NEAR-FIELD AUTOMOTIVE SAR
8035	Identifying Attack-Specific Signatures in Adversarial Examples
2926	IFNET: IMAGING AND FOCUSING NETWORK FOR HANDHELD MMWAVE DEVICES
3369	IFNET: INTEGRATING DATA AUGMENTATION AND DECOUPLED ATTENTION FUSION FOR 3D OBJECT DETECTION

<b>Paper Number</b>	<b>Paper Title</b>
4395	IHT-Inspired Neural Network for Single-Snapshot DOA Estimation with Sparse Linear Arrays
1827	Image Aesthetics Assessment via Learnable Queries
4010	Image Attribution by Generating Images
4635	IMAGE AUGMENTATION WITH CONTROLLED DIFFUSION FOR WEAKLY-SUPERVISED SEMANTIC SEGMENTATION
5591	Image Coding for Analytics via Adversarially Augmented Adaptation
4617	IMAGE HARMONIZATION Based on Hierarchical Dynamics
2343	Image Mixing and Gradient Smoothing to Enhance the SAR Image Attack Transferability
7189	IMAGE RESTORATION WITH GENERALIZED L2 LOSS AND CONVERGENT PLUG-AND-PLAY PRIOR
3919	IMAGE RETRIEVAL WITH COMPOSED QUERY BY MULTI-SCALE MULTI-MODAL FUSION
8629	IMAGE STEGANOGRAPHY WITH DEEP ORTHOGONAL FUSION OF MULTI-SCALE CHANNEL ATTENTION
3330	IMAGE2POINTS: A 3D POINT-BASED CONTEXT CLUSTERS GAN FOR HIGH-QUALITY PET IMAGE RECONSTRUCTION
4838	IMAGING AN EVOLVING BLACK HOLE BY LEVERAGING SHARED STRUCTURE
1151	IMFIT: Normal Estimation Via Learning Neural Implicit Surface
6602	IMITATING THE HUMAN VISUAL SYSTEM FOR SCANPATH PREDICTING
9486	IMPACT OF SAMPLING STRATEGIES ON THE MONITORING OF CLIMATE REGIME SHIFTS WITH A LEARNING DATA ASSIMILATION METHOD
1698	IMPLICIT ENHANCEMENT OF TARGET SPEAKER IN SPEAKER-ADAPTIVE ASR THROUGH EFFICIENT JOINT OPTIMIZATION
2472	IMPLICIT FOREGROUND-GUIDED NETWORK FOR ANOMALY DETECTION AND LOCALIZATION
3752	Implicit Neural Multiple Description for DNA-based data storage

<b>Paper Number</b>	<b>Paper Title</b>
1893	IMPLICIT NEURAL REPRESENTATION FOR LOW-OVERHEAD GRAPH-BASED HOLOGRAPHIC-TYPE COMMUNICATIONS
6058	Implicit-Knowledge-Guided Align before Understanding for KB-VQA
9370	Importance of negative sampling in weak label learning
9508	IMPORTANCE SAMPLING BASED FEDERATED UNSUPERVISED REPRESENTATION LEARNING
7404	IMPOSING EARLY AND ASYMPTOTIC CONSTRAINTS ON LIGME WITH APPLICATION TO NONCONVEX ENHANCEMENT OF FUSED LASSO MODELS
6837	IMPROVE DEEP FOREST WITH LEARNABLE LAYERWISE AUGMENTATION POLICY SCHEDULES
9991	IMPROVED CHILDREN'S AUTOMATIC SPEECH RECOGNITION COMBINING ADAPTERS AND SYNTHETIC DATA AUGMENTATION
9473	IMPROVED IMAGE CAPTIONING VIA KNOWLEDGE GRAPH-AUGMENTED MODELS
6066	IMPROVED SCREEN CONTENT CODING IN VVC USING SOFT CONTEXT FORMATION
2569	Improving acoustic echo cancellation by exploring speech and echo affinity with multi-head attention
4396	IMPROVING ACOUSTIC ECHO CANCELLATION FOR VOICE ASSISTANTS USING NEURAL ECHO SUPPRESSION AND MULTI-MICROPHONE NOISE REDUCTION
8039	IMPROVING ASR CONTEXTUAL BIASING WITH GUIDED ATTENTION
3351	IMPROVING ATTENTION-BASED END-TO-END SPEECH RECOGNITION BY MONOTONIC ALIGNMENT ATTENTION MATRIX RECONSTRUCTION
1863	IMPROVING AUDIO CAPTIONING MODELS WITH FINE-GRAINED AUDIO FEATURES, TEXT EMBEDDING SUPERVISION, AND LLM MIX-UP AUGMENTATION
6695	IMPROVING BIOMEDICAL ENTITY LINKING WITH RETRIEVAL-ENHANCED LEARNING
7261	IMPROVING CHINESE SPELLING CORRECTION WITH TEXT-PHONETICS DIFFERENTIATION AND ADAPTIVE FUSION

<b>Paper Number</b>	<b>Paper Title</b>
8142	IMPROVING CONTINUAL LEARNING OF ACOUSTIC SCENE CLASSIFICATION VIA MUTUAL INFORMATION OPTIMIZATION
1966	Improving Cross-domain Few-shot Classification with Multilayer Perceptron
11883	IMPROVING DATA-DRIVEN RF SIGNAL SEPARATION WITH SOI-MATCHED AUTOENCODERS
4011	IMPROVING DESIGN OF INPUT CONDITION INVARIANT SPEECH ENHANCEMENT
7059	IMPROVING DOMAIN GENERALIZATION IN SPEECH EMOTION RECOGNITION WITH WHISPER
4015	IMPROVING KINYARWANDA SPEECH RECOGNITION VIA SEMI-SUPERVISED LEARNING
9662	IMPROVING LANGUAGE MODEL-BASED ZERO-SHOT TEXT-TO-SPEECH SYNTHESIS WITH MULTI-SCALE ACOUSTIC PROMPTS
2067	Improving Learned Video Compression by Exploring Spatial Redundancy
7211	IMPROVING LIMITED SUPERVISED FOOT ULCER SEGMENTATION USING CROSS-DOMAIN AUGMENTATION STRATEGIES
7326	Improving Long Text Understanding with Knowledge Distilled from Summarization Model
7408	IMPROVING MEDICAL DIALOGUE GENERATION WITH ABSTRACT MEANING REPRESENTATIONS
2196	Improving Motion Deblur by Multi-Output Learning
7305	IMPROVING MULTI-MODAL EMOTION RECOGNITION USING ENTROPY-BASED FUSION AND PRUNING-BASED NETWORK ARCHITECTURE OPTIMIZATION
8889	IMPROVING MULTI-SPEAKER ASR WITH OVERLAP-AWARE ENCODING AND MONOTONIC ATTENTION
2760	IMPROVING MUSIC SOURCE SEPARATION WITH SIMO STEREO BAND-SPLIT RNN
7652	IMPROVING NEURAL DIARIZATION THROUGH SPEAKER ATTRIBUTE ATTRACTORS AND LOCAL DEPENDENCY MODELING

<b>Paper Number</b>	<b>Paper Title</b>
4656	IMPROVING OPEN-SET RECOGNITION WITH BAYESIAN METRIC LEARNING
4262	Improving Oral Reading Fluency Assessment through Sub-sequence Matching of Acoustic Word Embeddings
8883	IMPROVING RADIOLOGY REPORT GENERATION WITH D <sup>2</sup> -NET: WHEN DIFFUSION MEETS DISCRIMINATOR
7063	Improving Short Utterance Anti-Spoofing with AASIST2
7169	Improving Speaker-independent Speech Emotion Recognition Using Dynamic Joint Distribution Adaptation
9093	IMPROVING SPEECH ATTENUATION IN HEADPHONES USING HARMONIC MODEL DECOMPOSITION AND MULTIPLE-FREQUENCY ANC
1865	Improving Speech Emotion Recognition with Unsupervised Speaking Style Transfer
8746	IMPROVING SPEECH RECOGNITION FOR AFRICAN AMERICAN ENGLISH WITH AUDIO CLASSIFICATION
7194	IMPROVING SPEED/ACCURACY TRADEOFF FOR ONLINE STREAMING ASR VIA REAL-VALUED AND TRAINABLE STRIDES
9208	IMPROVING TARGET SOUND EXTRACTION WITH TIMESTAMP KNOWLEDGE DISTILLATION
1359	IMPROVING VGG-STYLE CONVNET FOR JPEG STEGANALYSIS
2795	IMPROVING VISION-INSPIRED KEYWORD SPOTTING USING DYNAMIC MODULE SKIPPING IN STREAMING CONFORMER ENCODER
2904	IMPROVING VISUAL QUALITY AND TRANSFERABILITY OF ADVERSARIAL ATTACKS ON FACE RECOGNITION SIMULTANEOUSLY WITH ADVERSARIAL RESTORATION
9022	INAPPROPRIATE PAUSE DETECTION IN DYSARTHIC SPEECH USING LARGE-SCALE SPEECH RECOGNITION
6555	INCOMPLETE MULTI-VIEW CLUSTERING VIA INFERENCE AND EVALUATION
8188	INCOMPLETE MULTI-VIEW REPRESENTATION LEARNING THROUGH ANCHOR GRAPH-BASED GCN AND INFORMATION BOTTLENECK

<b>Paper Number</b>	<b>Paper Title</b>
1678	Incomplete Observations Bias Suppression for Abductive Natural Language Inference
1612	In-Context Learning for Few-Shot Nested Named Entity Recognition
2147	IN-CONTEXT PROMPT EDITING FOR CONDITIONAL AUDIO GENERATION
8461	INCPROMPT: TASK-AWARE INCREMENTAL PROMPTING FOR REHEARSAL-FREE CLASS-INCREMENTAL LEARNING
6483	Incremental Tensor Decomposition for Few Shot Neural Radiance Field
7917	Inducing Inductive Bias in Vision Transformer for EEG Classification
5879	INFERENCE OF GENETIC EFFECTS VIA APPROXIMATE MESSAGE PASSING
7637	INFERENCE OF TIME-VARYING GRAPH TOPOLOGIES VIA GAUSSIAN PROCESSES
6212	INFERRING THE GRAPH OF NETWORKED DYNAMICAL SYSTEMS UNDER PARTIAL OBSERVABILITY AND SPATIALLY COLORED NOISE
9401	Inferring Time Varying Signals Over Uncertain Graphs
7939	INNOVATIVE METHODS FOR NON-DESTRUCTIVE INSPECTION OF HANDWRITTEN DOCUMENTS
2515	INPUTMIX: A STRATEGY TO REGULARIZE AND BALANCE MULTI-MODALITY AND MULTI-VIEW MODEL LEARNING
2529	Instant Photorealistic Neural Radiance Fields Stylization
8705	INTEGRATED LOCALIZATION AND COMMUNICATION IN 3GPP INDUSTRIAL ENVIRONMENTS
2941	INTEGRATED SENSING AND COMMUNICATION IN UNLICENSED MMWAVE BANDS: JOINT BEAMFORMING TRAINING AND ENERGY ALLOCATION
7004	INTEGRATING LANGUAGE MODELS WITH SYMBOLIC FORMULAS FOR FIRST-ORDER LOGIC REASONING
2352	INTEGRATING SENSING, COMMUNICATION, AND COMPUTATION IN THE SKY



<b>Paper Number</b>	<b>Paper Title</b>
5790	INTELLIGENT CARDIAC AUSCULTATION FOR MURMUR DETECTION VIA PARALLEL-ATTENTIVE MODELS WITH UNCERTAINTY ESTIMATION
11532	INTER-FREQUENCY PHASE DIFFERENCE FOR PHASE RECONSTRUCTION USING DEEP NEURAL NETWORKS AND MAXIMUM LIKELIHOOD
8378	INTER-MODALITY AND INTRA-SAMPLE ALIGNMENT FOR MULTI-MODAL EMOTION RECOGNITION
4760	INTERNAL LOCATION ASSISTANCE FOR TEMPORAL ACTION PROPOSAL GENERATION
1971	INTERPRETABLE FACE AGING: ENHANCING CONDITIONAL ADVERSARIAL AUTOENCODERS WITH LIME EXPLANATIONS
6915	INTERPRETABLE MULTIMODAL OUT-OF-CONTEXT DETECTION WITH SOFT LOGIC REGULARIZATION
9518	INTERPRETABLE POLICY EXTRACTION WITH NEURO-SYMBOLIC REINFORCEMENT LEARNING
2784	INTERPRETING MEMORIZATION IN DEEP LEARNING FROM DATA DISTRIBUTION
11467	INTERPRETING THE CONTRIBUTION OF SENSORS IN BLIND SOURCE EXTRACTION BY MEANS OF SHAPLEY VALUES
3020	In-the-Wild Physiological-based Stress Detection Using Federated Strategy
1846	INTRODUCING MULTILINGUAL PHONETIC INFORMATION TO SPEAKER EMBEDDING FOR SPEAKER VERIFICATION
5589	INVARIANT MOTION REPRESENTATION LEARNING FOR 3D TALKING FACE SYNTHESIS
8975	INVARIANTOODG: LEARNING INVARIANT FEATURES OF POINT CLOUDS FOR OUT-OF-DISTRIBUTION GENERALIZATION
11460	INVERSE IMAGE FREQUENCY FOR LONG-TAILED IMAGE RECOGNITION
2689	Inversive-Reasoning Augmentation for Natural Language Inference
6433	INVERTEDFONTNET: FONT WATERMARKING BASED ON PERTURBING STYLE MANIFOLD

<b>Paper Number</b>	<b>Paper Title</b>
2600	Invertible Mosaic Image Hiding Network for Very Large Capacity Image Steganography
1637	INVERTIBLE VOICE CONVERSION WITH PARALLEL DATA
8018	Investigating End-to-end ASR Architectures for Long form Audio Transcription
7444	INVESTIGATING PERSONALIZATION METHODS IN TEXT TO MUSIC GENERATION
4982	Investigating salient representations and label Variance in Dimensional Speech Emotion Analysis
7993	INVESTIGATING SELF-SUPERVISED DEEP REPRESENTATIONS FOR EEG-BASED AUDITORY ATTENTION DECODING
5301	INVESTIGATING THE CLUSTERS DISCOVERED BY PRE-TRAINED AV-HUBERT
4922	IPCL: ITERATIVE PSEUDO-SUPERVISED CONTRASTIVE LEARNING TO IMPROVE SELF-SUPERVISED FEATURE REPRESENTATION
9583	IPHONMATCHNET: ZERO-SHOT USER-DEFINED KEYWORD SPOTTING USING IMPLICIT ACOUSTIC ECHO CANCELLATION
2825	IRLSG: INVARIANT REPRESENTATION LEARNING FOR SINGLE-DOMAIN GENERALIZATION IN MEDICAL IMAGE SEGMENTATION
8496	IRREGULARITY-AWARE BANDLIMITED APPROXIMATION FOR GRAPH SIGNAL INTERPOLATION
1282	IRS-Assisted Covert Communication with a BPP Distributed Warden outside a Safety Zone
8468	IRS-Assisted Joint Sensing and Communication Design for Autonomous Driving
4479	ISAC Beamforming Optimization for Robust Transmission in Dynamic mmWave MIMO Networks
5396	ITERATIVE AUTOREGRESSIVE GENERATION FOR ABSTRACTIVE SUMMARIZATION
4418	ITERATIVELY PRECONDITIONED GUIDANCE OF DENOISING (DIFFUSION) MODELS FOR IMAGE RESTORATION

<b>Paper Number</b>	<b>Paper Title</b>
1212	J-MAE: JIGSAW MEETS MASKED AUTOENCODERS IN X-RAY SECURITY INSPECTION
2626	JM-CLIP: A Joint Modal Similarity Contrastive Learning Model for Video-Text Retrieval
7225	JOINT ADMISSION CONTROL AND BEAMFORMER DESIGN FOR MOBILE USERS: STAY HERE OR MOVE TO A BETTER POSITION?
7312	JOINT BEAMFORMING AND COMPRESSION DESIGN FOR PER-ANTENNA POWER CONSTRAINED COOPERATIVE CELLULAR NETWORKS
8507	Joint Blind Deconvolution and Demixing of Sparse Signals via Factorization and Nonconvex Optimization
7759	JOINT CHANNEL ESTIMATION AND DATA DETECTION IN MASSIVE MIMO SYSTEMS BASED ON DIFFUSION MODELS
2380	JOINT CLASSIFICATION OF HYPERSPECTRAL AND LIDAR DATA USING CROSS-MODAL HIERARCHICAL FREQUENCY FUSION NETWORK
7968	Joint Computing and Communication Resource Allocation for TDMA-Based Binary Computation Offloading
8084	JOINT DEMOSAICING AND DENOISING WITH DOUBLE DEEP IMAGE PRIORS
11542	JOINT DEREVERBERATION AND BEAMFORMING WITH BLIND ESTIMATION OF THE SHAPE PARAMETER OF THE DESIRED SOURCE PRIOR
1784	JOINT DOA ESTIMATION AND DISTORTED SENSOR DETECTION UNDER ENTANGLED LOW-RANK AND ROW-SPARSE CONSTRAINTS
5687	Joint Embedding Learning and Latent Subspace Probing for Cross-domain Few-shot Keyword Spotting
4687	Joint End-to-End Spoken Language Understanding and Automatic Speech Recognition Training based on Unified Speech-to-Text Pre-training
2791	JOINT INDSCAL DECOMPOSITION MEETS BLIND SOURCE SEPARATION
4771	JOINT INFERENCE OF SPEAKER DIARIZATION AND ASR WITH MULTI-STAGE INFORMATION SHARING

<b>Paper Number</b>	<b>Paper Title</b>
1105	JOINT LEARNING OF IDENTITY AND VEIN FEATURES FOR ENHANCED REPRESENTATIONS IN VASCULAR BIOMETRICS
1531	JOINT MULTI-BAND DOA ESTIMATION USING LOW-RANK MATRIX RECOVERY
2548	JOINT MULTI-FACTS REASONING NETWORK FOR COMPLEX TEMPORAL QUESTION ANSWERING OVER KNOWLEDGE GRAPH
7566	JOINT MUSIC AND LANGUAGE ATTENTION MODELS FOR ZERO-SHOT MUSIC TAGGING
7778	JOINT NEAR-FIELD TARGET TRACKING AND COMMUNICATIONS WITH FULL DUPLEX HOLOGRAPHIC MIMO
7919	Joint Ranging and Phase Offset Estimation of Multiple Aviation Vehicles using Secondary Radar
9797	Joint Robust Optimal Transmit and Receive Beamforming Designs for a DFRC System for the MIMO Radar and Secondary Multicast Communication in a Cognitive Radio Network
11522	JOINT SEPARATION AND LOCALIZATION OF MOVING SOUND SOURCES BASED ON NEURAL FULL-RANK SPATIAL COVARIANCE ANALYSIS
7028	JOINT SIGNAL INTERPOLATION / TIME-VARYING GRAPH ESTIMATION VIA SMOOTHNESS AND LOW-RANK PRIORS
9651	Joint Signal Recovery and Graph Learning from Incomplete Time-Series
9027	JOINT SPATIO-TEMPORAL FILTERING OF MOTION IMAGERY EEG SIGNALS FOR DATA ALIGNMENT IN TRANSFER LEARNING
7910	JOINT TRANSMIT PRECODERS AND PASSIVE REFLECTION BEAMFORMER DESIGN IN IRS-AIDED IOT NETWORKS
4593	JOINT UNSUPERVISED AND SUPERVISED TRAINING FOR AUTOMATIC SPEECH RECOGNITION VIA BILEVEL OPTIMIZATION
6840	JOINTLY LEARNING SELECTION MATRICES FOR TRANSMITTERS, RECEIVERS AND FOURIER COEFFICIENTS IN MULTICHANNEL IMAGING

<b>Paper Number</b>	<b>Paper Title</b>
1958	JOINT-SEMANTICS MULTI-SIMILARITY HASHING FOR CROSS-MODAL RETRIEVAL
3208	JPEG ENCRYPTION WITH DC PREDICTION AND RUN-BASED RS PAIRS PERMUTATION
3050	JPIS: A JOINT MODEL FOR PROFILE-BASED INTENT DETECTION AND SLOT FILLING WITH SLOT-TO-INTENT ATTENTION
7262	KALMAN FILTER FOR TRACKING NETWORK DYNAMIC
8698	Kalman Filtering with Unlimited Sensing
7438	KC-Prompt: End-to-end Knowledge-Complementary Prompting for Rehearsal-free Continual Learning
5015	KD-Former: Transformer Knowledge Distillation for Image Matting
4464	KEEP DECODING PARALLEL WITH EFFECTIVE KNOWLEDGE DISTILLATION FROM LANGUAGE MODELS TO END-TO-END SPEECH RECOGNISERS
8472	KEEP KNOWLEDGE IN PERCEPTION: ZERO-SHOT IMAGE AESTHETIC ASSESSMENT
7734	KENET:KNOWLEDGE-ENHANCED DOC-LABEL ATTENTION NETWORK FOR MULTI-LABEL TEXT CLASSIFICATION
10125	KEY POINTS CENTERED SPARSE HASHING FOR CROSS-MODAL RETRIEVAL
8970	Killing it with Zero-Shot: Adversarially Robust Novelty Detection
8299	K-Means Clustering based on Chebyshev Polynomial Graph Filtering
4747	KNN-CTC: ENHANCING ASR VIA RETRIEVAL OF CTC PSEUDO LABELS
3311	KNOWLEDGE-AWARE PROMPT LEARNING FRAMEWORK FOR KOREAN-CHINESE MICROBLOG SENTIMENT ANALYSIS
7941	KNOWLEDGE-BASED CONVOLUTIONAL NEURAL NETWORK FOR THE SIMULATION AND PREDICTION OF TWO-PHASE DARCY FLOWS
11600	Kronecker-Product Beamforming with Sparse Concentric Circular Arrays

<b>Paper Number</b>	<b>Paper Title</b>
11877	KS-NET: MULTI-BAND JOINT SPEECH RESTORATION AND ENHANCEMENT NETWORK FOR 2024 ICASSP SSI CHALLENGE
9981	L1-aware Multilingual Mispronunciation Detection Framework
6322	LABCLIP: LABEL-ENHANCED CLIP FOR IMPROVING ZERO-SHOT TEXT CLASSIFICATION
5192	LABEL CORRECTION FOR SKETCH-BASED 3D SHAPE RETRIEVAL
5355	LABEL DEPENDENCIES-AWARE SET PREDICTION NETWORKS FOR MULTI-LABEL TEXT CLASSIFICATION
5113	LABEL RECTIFIED AND GRAPH ADAPTIVE SEMI-SUPERVISED REGRESSION FOR ELECTRODE SHIFTED GESTURE RECOGNITION
7068	LABEL-AWARE AUXILIARY LEARNING FOR DIALOGUE STATE TRACKING
2006	LACVIT: A LABEL-AWARE CONTRASTIVE FINE-TUNING FRAMEWORK FOR VISION TRANSFORMERS
9803	LANGUAGE GUIDED ADVERSARIAL PURIFICATION
1299	Language Model is a Branch Predictor for Simultaneous Machine Translation
4012	LANGUAGE-DRIVEN OPEN-VOCABULARY 3D SEMANTIC SEGMENTATION WITH KNOWLEDGE DISTILLATION
2339	Language-Driven Ordinal Learning for Imbalanced Head Pose Estimation
2258	LANGUAGE-FREE COMPOSITIONAL ACTION GENERATION VIA DECOUPLING REFINEMENT
1288	Language-guided Few-shot Semantic Segmentation
9515	LANGUAGE-ORIENTED COMMUNICATION WITH SEMANTIC CODING AND KNOWLEDGE DISTILLATION FOR TEXT-TO-IMAGE GENERATION
3746	LANGWAVE: REALISTIC VOICE GENERATION BASED ON HIGH-ORDER LANGEVIN DYNAMICS
7156	LARGE COVARIANCE MATRIX ESTIMATION BASED ON FACTOR MODELS VIA NONCONVEX OPTIMIZATION

<b>Paper Number</b>	<b>Paper Title</b>
4795	LARGE LANGUAGE MODEL-BASED EMOTIONAL SPEECH ANNOTATION USING CONTEXT AND ACOUSTIC FEATURE FOR SPEECH EMOTION RECOGNITION
4406	LARGE LANGUAGE MODELS AS A PROXY FOR HUMAN EVALUATION IN ASSESSING THE COMPREHENSIBILITY OF DISORDERED SPEECH TRANSCRIPTION
2883	LARGE LANGUAGE MODELS AUGMENTED RATING PREDICTION IN RECOMMENDER SYSTEM
1629	Large Scale Self-Supervised Pretraining for Active Speaker Detection
4690	LARGE-SCALE MULTI-VIEW MULTIPLE CLUSTERING
3317	Latent Degradation Representation Constraint for Single Image Deraining
5161	LATENT FILLING: LATENT SPACE DATA AUGMENTATION FOR ZERO-SHOT SPEECH SYNTHESIS
3582	LCB-NET: LONG-CONTEXT BIASING FOR AUDIO-VISUAL SPEECH RECOGNITION
7585	LEAKY WAVEGUIDE ANTENNAS FOR DOWNLINK WIDEBAND THZ COMMUNICATIONS
9222	LEARN FROM ZOOM: DECOUPLED SUPERVISED CONTRASTIVE LEARNING FOR WCE IMAGE CLASSIFICATION
1724	LEARN TO CLUSTER FACES WITH BETTER SUBGRAPHS
6123	LEARN TO TRACK-BEFORE-DETECT VIA NEURAL DYNAMIC PROGRAMMING
4264	Learnable Statistical Moments Pooling for Automatic Modulation Classification
1402	LEARNED ISTA WITH ERROR-BASED THRESHOLDING FOR ADAPTIVE SPARSE CODING
4102	LEARNED LAYERED CODING FOR SUCCESSIVE REFINEMENT IN THE WYNER-ZIV PROBLEM
6383	LEARNED VIDEO COMPRESSION WITH SPATIAL-TEMPORAL OPTIMIZATION
7343	LEARNING A CONVEX PATCH-BASED SYNTHESIS MODEL VIA DEEP EQUILIBRIUM

<b>Paper Number</b>	<b>Paper Title</b>
3794	LEARNING A LOW-RANK FEATURE REPRESENTATION: ACHIEVING BETTER TRADE-OFF BETWEEN STABILITY AND PLASTICITY IN CONTINUAL LEARNING
2166	Learning Active Subspaces for Effective and Scalable Uncertainty Quantification in Deep Neural Networks
8114	LEARNING AROUSAL-VALENCE REPRESENTATION FROM CATEGORICAL EMOTION LABELS OF SPEECH
2143	LEARNING AUDIO CONCEPTS FROM COUNTERFACTUAL NATURAL LANGUAGE
8656	LEARNING CONTEXTUALIZED REPRESENTATION ON DISCRETE SPACE VIA HIERARCHICAL PRODUCT QUANTIZATION
9339	Learning Density Regulated and Multi-view Consistent Unsigned Distance Fields
6056	LEARNING DISCRIMINATIVE STYLE REPRESENTATIONS FOR UNSUPERVISED AND FEW-SHOT ARTISTIC PORTRAIT DRAWING GENERATION
8240	LEARNING DISENTANGLED SPEECH REPRESENTATIONS WITH CONTRASTIVE LEARNING AND TIME-INVARIANT RETRIEVAL
4294	LEARNING DYNAMICS OF LOW-PRECISION CLIPPED SGD WITH MOMENTUM
3543	LEARNING EMOTION-INVARIANT SPEAKER REPRESENTATIONS FOR SPEAKER VERIFICATION
8307	Learning Fine-Grained Information Alignment for Calibrated Cross-Modal Retrieval
3840	LEARNING FROM EASY TO HARD: MULTI-TASK LEARNING WITH DATA SCHEDULING
4494	LEARNING FROM TAXONOMY: MULTI-LABEL FEW-SHOT CLASSIFICATION FOR EVERYDAY SOUND RECOGNITION
2357	LEARNING GENERALIZABLE VISUAL REPRESENTATIONS VIA SELF-SUPERVISED INFORMATION BOTTLENECK
9095	Learning Graphs and Simplicial Complexes from Data
1508	LEARNING HYBRID NEGATIVE PROBABILITY MODEL FOR WEAKLY-SUPERVISED WHOLE SLIDE IMAGE RECOGNITION



<b>Paper Number</b>	<b>Paper Title</b>
1438	Learning Inference-Time Drift Sensor-Actuator for Domain Generalization
1507	Learning Invariant Representation with Consistency and Diversity for Semi-supervised Source Hypothesis Transfer
2756	LEARNING MULTIPLEX GRAPH WITH INTER-LAYER COUPLING
1549	LEARNING MULTISCALE CONSISTENCY FOR SELF-SUPERVISED ELECTRON MICROSCOPY INSTANCE SEGMENTATION
7459	LEARNING ONTOLOGY INFORMED REPRESENTATIONS WITH CONSTRAINTS FOR ACOUSTIC EVENT DETECTION
5635	LEARNING REPRESENTATIONS FROM EXPLAINABLE AND CONNECTIONIST APPROACHES FOR VISUAL QUESTION ANSWERING
8759	LEARNING SEMANTIC INFORMATION FROM RAW AUDIO SIGNAL USING BOTH CONTEXTUAL AND PHONETIC REPRESENTATIONS
7350	LEARNING SIGNALS AND GRAPHS FROM TIME-SERIES GRAPH DATA WITH FEW CAUSES
8359	LEARNING SPATIO-TEMPORAL RELATIONS WITH MULTI-SCALE INTEGRATED PERCEPTION FOR VIDEO ANOMALY DETECTION
7716	LEARNING SPEAKER-LISTENER MUTUAL HEAD ORIENTATION BY LEVERAGING HRTF AND VOICE DIRECTIVITY ON HEADPHONES
9615	LEARNING SPECTRAL CANONICAL F-CORRELATION REPRESENTATION FOR FACE SUPER-RESOLUTION
2104	Learning Speech Representation From Contrastive Token-Acoustic Pretraining
11562	LEARNING STOCHASTIC GRAPH NEURAL NETWORKS WITH CONSTRAINED VARIANCE
9807	Learning the Barankin Lower Bound on DOA estimation error
11450	Learning to Bound: A Generative Cramér-Rao Bound
3508	Learning with Non-Uniform Label Noise: A Cluster-Dependent Weakly Supervised Approach

<b>Paper Number</b>	<b>Paper Title</b>
5586	Least-Effort Adversarial Attack Against Gait-based Identity Recognition System
3173	LEFORMER: A HYBRID CNN-TRANSFORMER ARCHITECTURE FOR ACCURATE LAKE EXTRACTION FROM REMOTE SENSING IMAGERY
8338	Lesion-aware Open Set Medical Image Recognition with Domain Shift
7412	LESS PEAKY AND MORE ACCURATE CTC FORCED ALIGNMENT BY LABEL PRIORS
1025	LEVERAGE CAUSAL GRAPHS AND RUMOR-REFUTING TEXTS FOR INTERPRETABLE RUMOR ANALYSIS
9265	LEVERAGING BIASES IN LARGE LANGUAGE MODELS: “BIAS-KNN” FOR EFFECTIVE FEW-SHOT LEARNING
9516	Leveraging Data Collection and Unsupervised Learning for Code-switched Tunisian Arabic Automatic Speech Recognition
11902	LEVERAGING EFFECTIVE LANGUAGE AND SPEAKER CONDITIONING IN INDIC TTS FOR LIMMITS 2024 CHALLENGE
4555	Leveraging In-the-Wild Data for Effective Self-Supervised Pretraining in Speaker Recognition
8369	LEVERAGING LARGE LANGUAGE MODELS FOR EXPLOITING ASR UNCERTAINTY
8336	Leveraging Large Pretrained Models for Line-by-Line Spoken Program Recognition
9964	LEVERAGING NOISY LABELS OF NEAREST NEIGHBORS FOR LABEL CORRECTION AND SAMPLE SELECTION
2849	Leveraging Redundancy in Feature for Efficient Learned Image Compression
4381	LEVERAGING SELF-SUPERVISED SPEECH REPRESENTATIONS FOR DOMAIN ADAPTATION IN SPEECH ENHANCEMENT
3956	LEVERAGING SOUND LOCALIZATION TO IMPROVE CONTINUOUS SPEAKER SEPARATION
5102	LEVERAGING SPEECH PTM, TEXT LLM, AND EMOTIONAL TTS FOR SPEECH EMOTION RECOGNITION

<b>Paper Number</b>	<b>Paper Title</b>
9730	LEVERAGING TENSOR SUBSPACE PRIOR: ENHANCED SUM OF NUCLEAR NORM MINIMIZATION FOR TENSOR COMPLETION
2764	LEVERAGING TIMESTAMP INFORMATION FOR SERIALIZED JOINT STREAMING RECOGNITION AND TRANSLATION
8785	Leveraging Visual Handicaps for Text-based Reinforcement Learning
4728	LIBRIHEAVY: A 50,000 HOURS ASR CORPUS WITH PUNCTUATION CASING AND CONTEXT
3738	LIGHTCODEC: A HIGH FIDELITY NEURAL AUDIO CODEC WITH LOW COMPUTATION COMPLEXITY
7039	Lighting Image/Video Style Transfer Methods by Iterative Channel Pruning
4613	Lightweight high-resolution Subject Matting in the Real World
3272	Lightweight Multi-Axial Transformer with Frequency Prompt for Single Channel Speech Enhancement
3916	LIKELIHOOD CONSENSUS 2.0: REDUCING INTERAGENT COMMUNICATION IN DISTRIBUTED BAYESIAN TARGET TRACKING
11898	LIMITS'24: MULTI-SPEAKER, MULTI-LINGUAL INDIC TTS WITH VOICE CLONING
10438	Linear Complexity Gibbs Sampling for Generalized Labeled Multi-Bernoulli Filtering
11466	Linearly-involved Moreau-Enhanced-over-Subspace Model: Debiased Sparse Modeling and Stable Outlier-Robust Regression
5853	LIPSCHITZ-CONSTRAINED CONVOLUTIONAL LAYERS USING CONVEX PROJECTION
2817	LITEVSR: EFFICIENT VISUAL SPEECH RECOGNITION BY LEARNING FROM SPEECH REPRESENTATIONS OF UNLABELED DATA
4047	LIVE ITERATIVE PTYCHOGRAPHY WITH PROJECTION-BASED ALGORITHMS
1712	LK-UNET: LARGE KERNEL DESIGN FOR 3D MEDICAL IMAGE SEGMENTATION

<b>Paper Number</b>	<b>Paper Title</b>
9699	LLET: LIGHTWEIGHT LEXICON-ENHANCED TRANSFORMER FOR CHINESE NER
2821	LOCAL AND GLOBAL FEATURE ADAPTIVE ADJUSTMENT NETWORK FOR REMOTE SENSING IMAGE SCENE CLASSIFICATION
6997	local and global: text matching via syntax graph calibration
9603	Local Contrast Prior-Guided Cross Aggregation Model for Effective Infrared Small Target Detection
1251	LOCAL DISTANCE CORRELATION EMBEDDING FOR TIME-SERIES ANALYSIS ON RIEMANNIAN MANIFOLDS
10353	LOCAL INFORMATION GUIDED GLOBAL INTEGRATION FOR INFRARED SMALL TARGET DETECTION
8005	LOCAL OPTIMIZATION NETWORKS FOR MULTI-VIEW MULTI-PERSON HUMAN POSTURE ESTIMATION
2119	Locality-Enhanced Transformer for Semantic Segmentation of High-Resolution Remote Sensing Images
7765	LOCALIZATION AND TRACKING OF GOLD NANOPARTICLES USING MMWAVE FMCW RADAR
1994	LOCALIZATION IN SENSOR NETWORKS USING DISTRIBUTED LOW-RANK MATRIX COMPLETION
1914	LOCALIZING ACOUSTIC ENERGY IN SOUND FIELD SYNTHESIS BY DIRECTIONALLY WEIGHTED EXTERIOR RADIATION SUPPRESSION
3184	LOCATION OPTIMIZATION FOR RIS AIDED MMWAVE DOWNLINK NETWORK
6896	LOCSELECT: TARGET SPEAKER LOCALIZATION WITH AN AUDITORY SELECTIVE HEARING MECHANISM
9924	LoFi User Scheduling for Multiuser MIMO Wireless Systems
6938	LOFT: LATENT SPACE OPTIMIZATION AND GENERATOR FINE-TUNING FOR DEFENDING AGAINST DEEPPKES
6326	LONG TERM MEMORY-ENHANCED VIA CAUSAL REASONING FOR TEXT-TO-VIDEO RETRIEVAL
7857	LONGITUDINAL MODELING OF DEPRESSION SHIFTS USING SPEECH AND LANGUAGE

<b>Paper Number</b>	<b>Paper Title</b>
3846	LONG-TERM ACTION ANTICIPATION BASED ON CONTEXTUAL ALIGNMENT
7951	LONG-TERM SOCIAL INTERACTION CONTEXT: THE KEY TO EGOCENTRIC ADDRESSEE DETECTION
2496	Look, Listen and Recognise: Character-Aware Audio-Visual Subtitling
8901	Loop Structure-Aware Learning for Fully Automated Pulmonary Fissure Completeness Assessment
4857	LOSS MASKING IS NOT NEEDED IN DECODER-ONLY TRANSFORMER FOR DISCRETE-TOKEN-BASED ASR
2670	LOSSY COMPRESSION OF ADJACENCY MATRICES BY GRAPH FILTER BANKS
7175	Low Bitrate Loss Resilience Scheme For a Speech Enhancing Neural Codec
11913	LOW DOSE CBCT DENOISING USING A 3D U-NET
4145	LOW OVERHEAD DMG SENSING FOR VITAL SIGNS DETECTION
2375	LOW REDUNDANT ATTENTION NETWORK FOR EFFICIENT IMAGE SUPER-RESOLUTION
9716	Low-Complexity GLRT Based Quickest Detection with Unknown Parameters
4390	LOW-COMPLEXITY VECTOR SOURCE CODING FOR DISCRETE LONG SEQUENCES WITH UNKNOWN DISTRIBUTIONS
4138	LOW-LATENCY SPEECH ENHANCEMENT VIA SPEECH TOKEN GENERATION
8822	LOW-LIGHT RAW IMAGE ENHANCEMENT ON A DATASET SUFFERING LIGHT EFFECTS
11504	LOW-PAPR OFDM WAVEFORM DESIGN FOR RADAR AND COMMUNICATION SYSTEMS
7318	LOW-RANK COMPLETION BASED NORMAL GUIDED LIDAR POINT CLOUD UP-SAMPLING
2519	LOW-RANK CONSTRAINED MULTICHANNEL SIGNAL DENOISING CONSIDERING CHANNEL-DEPENDENT SENSITIVITY INSPIRED BY SELF-SUPERVISED LEARNING FOR OPTICAL FIBER SENSING

<b>Paper Number</b>	<b>Paper Title</b>
2398	LVC-LGMC: JOINT LOCAL AND GLOBAL MOTION COMPENSATION FOR LEARNED VIDEO COMPRESSION
10019	LV-SEGFORMER: TOWARDS MORE ACCURATE LEAF-VEIN SEGMENTATION WITH TRANSFORMER
8485	M <sup>3</sup> ARL: Moment-Embedded Mean-Field Multi-Agent Reinforcement Learning for Continuous Action Space
1355	M <sup>3</sup> TN: Multi-gate Mixture-of-Experts based Multi-valued Treatment Network for More Uplift Modeling
7579	M2BART: MULTILINGUAL AND MULTIMODAL ENCODER-DECODER PRE-TRAINING FOR ANY-TO-ANY MACHINE TRANSLATION
4499	M2SUM: MULTI-GRANULARITY SCALE-ADAPTIVE VIDEO SUMMARIZER TOWARDS INFORMATIVE CONTEXT REPRESENTATION LEARNING
6683	M3DSYNTH: A DATASET OF MEDICAL 3D IMAGES WITH AI-GENERATED LOCAL MANIPULATIONS
8731	M3SUM: A NOVEL UNSUPERVISED LANGUAGE-GUIDED VIDEO SUMMARIZATION
1843	M3TQA: MULTI-VIEW, MULTI-HOP AND MULTI-STAGE REASONING FOR TEMPORAL QUESTION ANSWERING
4194	MACCN: MULTI-MODAL ADAPTIVE CO-ATTENTION FUSION CONTRASTIVE LEARNING NETWORKS FOR FAKE NEWS DETECTION
3404	MaDE: MULTI-SCALE DECISION ENHANCEMENT FOR MULTI-AGENT REINFORCEMENT LEARNING
3571	MADRL-BASED UAVS TRAJECTORY DESIGN WITH ANTI-COLLISION MECHANISM IN VEHICULAR NETWORKS
3155	MAINLOBE DECEPTIVE JAMMER SUPPRESSION USING FDA-MIMO RADAR IN THE PRESENCE OF MULTIPATH PROPAGATION
8482	MAML-BASED 24-HOUR PERSONALIZED BLOOD PRESSURE ESTIMATION FROM WRIST PHOTOPLETHYSMOGRAPHY SIGNALS IN FREE-LIVING CONTEXT
6000	MANTICORE: AN UNSUPERVISED INTRUSION DETECTION SYSTEM BASED ON CONTRASTIVE LEARNING IN 5G NETWORKS

<b>Paper Number</b>	<b>Paper Title</b>
3990	MAPACHE: MASKED PARALLEL TRANSFORMER FOR ADVANCED SPEECH EDITING AND SYNTHESIS
3886	MAPFLOW: MULTI-AGENT PEDESTRIAN TRAJECTORY PREDICTION USING NORMALIZING FLOW
5076	Mask6D: Masked Pose Priors For 6D Object Pose Estimation
4210	MaskMark: Robust Neural Watermarking for Real and Synthetic Speech
9313	MaskSTR: Guide Scene Text Recognition Models with Masking
8481	MAS-NET: MIXED-FEATURE ATTENTION SIAMESE NETWORK FOR CHANGE DETECTION ON REMOTE SENSING IMAGES
6040	MATCHA-TTS: A FAST TTS ARCHITECTURE WITH CONDITIONAL FLOW MATCHING
2557	MATPR-UNET: A MULTI ATTENTION TWO-PATH RESIDUAL UNET FOR FOCAL CORTICAL DYSPLASIA LESIONS SEGMENTATION
4336	MATRIX FACTORIZATION IN TROPICAL AND MIXED TROPICAL-LINEAR ALGEBRAS
7388	MAX-AST: COMBINING CONVOLUTION, LOCAL AND GLOBAL SELF-ATTENTIONS FOR AUDIO EVENT CLASSIFICATION
2606	MAXIMAL CODING RATE REDUCTION FOR GRAPH EMBEDDINGS
11551	MAXIMUM LIKELIHOOD-BASED GRIDLESS DOA ESTIMATION USING STRUCTURED COVARIANCE MATRIX RECOVERY AND SBL WITH GRID REFINEMENT
4372	MAXIMUM-ENTROPY ADVERSARIAL AUDIO AUGMENTATION FOR KEYWORD SPOTTING
8362	MAX-MARGIN TRANSDUCER LOSS: IMPROVING SEQUENCE-DISCRIMINATIVE TRAINING USING A LARGE-MARGIN LEARNING STRATEGY
8257	Max-min Beamforming for Multi-User Massive MIMO Systems: An Alternating Projection-Based Approach

<b>Paper Number</b>	<b>Paper Title</b>
7720	MCM-CSD: MULTI-GRANULARITY CONTEXT MODELING WITH CONTRASTIVE SPEAKER DETECTION FOR EMOTION RECOGNITION IN REAL-TIME CONVERSATION
5311	MDAVIF: A MULTI-DOMAIN ACOUSTICAL-VISUAL INFORMATION FUSION MODEL FOR DEPRESSION RECOGNITION FROM VLOG DATA
5180	MDRT: MULTI-DOMAIN SYNTHETIC SPEECH LOCALIZATION
4402	MDX-GAN: ENHANCING PERCEPTUAL QUALITY IN MULTI-CLASS SOURCE SEPARATION VIA ADVERSARIAL TRAINING
2867	MEAT: Median-Ensemble Adversarial Training for Improving Robustness and Generalization
5054	MEDICAL VISION-LANGUAGE REPRESENTATION LEARNING WITH CROSS-MODAL MULTI-TEACHER CONTRASTIVE DISTILLATION
9717	MELS-TTS : MULTI-EMOTION MULTI-LINGUAL MULTI-SPEAKER TEXT-TO-SPEECH SYSTEM VIA DISENTANGLED STYLE TOKENS
8383	MEMORY EFFICIENT CORNER DETECTION FOR EVENT-DRIVEN DYNAMIC VISION SENSORS
8783	MEMORY SELF-CALIBRATED NETWORK FOR VISUAL GROUNDING
2630	Memory-augmented Dual-domain Unfolding Network for MRI reconstruction
6496	MEMORY-AUGMENTED ONLINE VIDEO ANOMALY DETECTION
9897	MEMORY-AUGMENTED SPEECH-TO-TEXT TRANSLATION WITH MULTI-SCALE CONTEXT TRANSLATION STRATEGY
7817	MEPE: A Minimalist Ensemble Policy Evaluation Operator for Deep Reinforcement Learning
2538	MERG: Multi-dimensional Edge Representation Generation Layer for Graph Neural Networks
3363	MERTECH: INSTRUMENT PLAYING TECHNIQUE DETECTION USING SELF-SUPERVISED PRETRAINED MODEL WITH MULTI-TASK FINETUNING
2935	MESH-RTUME: UNIVERSAL MANIFOLD EMBEDDING FOR ESTIMATING 3D RIGID TRANSFORMATIONS OF SURFACES



<b>Paper Number</b>	<b>Paper Title</b>
5844	META REPRESENTATION LEARNING METHOD FOR ROBUST SPEAKER VERIFICATION IN UNSEEN DOMAINS
1758	META STRUCTURE SEARCH FOR LINK WEIGHT PREDICTION IN HETEROGENEOUS GRAPH
4169	META-AF ECHO CANCELLATION FOR IMPROVED KEYWORD SPOTTING
2676	META-KNOWLEDGE ENHANCED DATA AUGMENTATION FOR FEDERATED PERSON RE-IDENTIFICATION
1852	META-LEARNING WITH VERSATILE LOSS GEOMETRIES FOR FAST ADAPTATION USING MIRROR DESCENT
8034	METASURFACE-BASED RECEIVERS WITH 1-BIT ADCS FOR MULTI-USER UPLINK COMMUNICATIONS
4899	MF-AED-AEC: SPEECH EMOTION RECOGNITION BY LEVERAGING MULTIMODAL FUSION, ASR ERROR DETECTION, AND ASR ERROR CORRECTION
2930	MFT-PCQA: Multi-modal Fusion Transformer for No-reference Point Cloud Quality Assessment
2194	MGRL: MUTUAL-GUIDANCE REPRESENTATION LEARNING FOR TEXT-TO-IMAGE PERSON RETRIEVAL
4943	MHPS: MULTIMODALITY-GUIDED HIERARCHICAL POLICY SEARCH FOR KNOWLEDGE GRAPH REASONING
2856	Micro-expression Recognition by Fusing Action Unit Detection and Spatio-temporal Features
9321	MICROPHONE CONVERSION: MITIGATING DEVICE VARIABILITY IN SOUND EVENT CLASSIFICATION
7493	MICROPHONE SUBSET SELECTION FOR THE WEIGHTED PREDICTION ERROR ALGORITHM USING A GROUP SPARSITY PENALTY
9539	MIDI-Voice: Expressive Zero-shot Singing Voice Synthesis via MIDI-driven Priors
4041	MIMO IMAGING METHOD WITH ITERATIVE-BASED SUPER-RESOLUTION FOR AUTOMOTIVE RADAR
2100	MINIMALLY-SUPERVISED SPEECH SYNTHESIS WITH CONDITIONAL DIFFUSION MODEL AND LANGUAGE MODEL: A COMPARATIVE STUDY OF SEMANTIC CODING

<b>Paper Number</b>	<b>Paper Title</b>
11497	MINIMIZING LOW-RANK MODELS OF HIGH-ORDER TENSORS: HARDNESS, SPAN, TIGHT RELAXATION, AND APPLICATIONS
8977	MIR-MLPOP: A MULTILINGUAL POP MUSIC DATASET WITH TIME-ALIGNED LYRICS AND AUDIO
5742	MISA: UNVEILING THE VULNERABILITIES IN SPLIT FEDERATED LEARNING
1921	MISSPECIFIED TIME-DELAY AND DOPPLER ESTIMATION OVER NON GAUSSIAN SCENARIOS
8441	MITIGATE REPLICATION AND COPYING IN DIFFUSION MODELS WITH GENERALIZED CAPTION AND DUAL FUSION ENHANCEMENT
7782	MITIGATING DATA INJECTION ATTACKS ON FEDERATED LEARNING
5127	MITIGATING INTRA-CLASS VARIANCE IN FEW-SHOT POINT CLOUD CLASSIFICATION
1462	MITIGATING OPTIMIZATION CONFLICT IN DOMAIN ADVERSARIAL NEURAL NETWORK VIA UNCERTAINTY-AWARE
3121	MIXED GRAPH SIGNAL ANALYSIS OF JOINT IMAGE DENOISING / INTERPOLATION
1104	MIXED INFORMED TRANSFORMER FOR FEW-SHOT MEDICAL IMAGE SEGMENTATION
4904	MIXED PRECISION NEURAL QUANTIZATION WITH MULTI-OBJECTIVE BAYESIAN OPTIMIZATION FOR ON-DEVICE DEPLOYMENT
8686	MIXED-ATTENTION AUTO ENCODER FOR MULTI-CLASS INDUSTRIAL ANOMALY DETECTION
6202	MLCA-AVSR: MULTI-LAYER CROSS ATTENTION FUSION BASED AUDIO-VISUAL SPEECH RECOGNITION
9248	MLMTD: A MULTI-LAYER MALICIOUS TRAFFIC DETECTION MODEL BASED ON MULTI-BRANCH OCTAVE CONVOLUTION AND ATTENTION MECHANISM
6042	MLPs Compass: What is learned when MLPs are combined with PLMs?

<b>Paper Number</b>	<b>Paper Title</b>
7784	MMAFLOW: MATCHING-GUIDED MOTION AGGREGATION FOR OPTICAL FLOW ESTIMATION
10205	MMBAT: A MULTI-TASK FRAMEWORK FOR MMWAVE-BASED HUMAN BODY RECONSTRUCTION AND TRANSLATION PREDICTION
4197	mmCount: Stationary Crowd Counting System Based on Commodity Millimeter-wave Radar
6777	MMHSV: A MULTIMODAL HANDWRITTEN SIGNATURE VERIFICATION FUSING DYNAMIC AND STATIC FEATURE
7247	MMRBN: RULE-BASED NETWORK FOR MULTIMODAL EMOTION RECOGNITION
6884	MMS: MORPHOLOGY-MIXUP STYLIZED DATA GENERATION FOR SINGLE DOMAIN GENERALIZATION IN MEDICAL IMAGE SEGMENTATION
8991	MODAL CONSENSUS AND CONTEXTUAL SEPARATION FOR WEAKLY SUPERVISED TEMPORAL ACTION LOCALIZATION
7898	MODALITY DROP-OUT FOR MULTIMODAL DEVICE DIRECTED SPEECH DETECTION USING VERBAL AND NON-VERBAL FEATURES
3042	MODALITY RE-BALANCE FOR VISUAL QUESTION ANSWERING: A CAUSAL FRAMEWORK
2614	Modality-dependent sentiments exploring for multi-modal sentiment classification
4765	MODEL-BASED LABEL-TO-IMAGE DIFFUSION FOR SEMI-SUPERVISED CHOROIDAL VESSEL SEGMENTATION
1397	MODEL-BASED LEARNING FOR LOCATION-TO-CHANNEL MAPPING
9758	MODELING INTRAPERSONAL AND INTERPERSONAL INFLUENCES FOR AUTOMATIC ESTIMATION OF THERAPIST EMPATHY IN COUNSELING CONVERSATION
7030	Modeling pseudo-speaker uncertainty in voice anonymization
7044	MODELING QUASI-PERIODIC DEPENDENCY VIA SELF-SUPERVISED PRE-TRAINING FOR RESPIRATORY SOUND CLASSIFICATION

<b>Paper Number</b>	<b>Paper Title</b>
2099	MODELING ROUTE REPRESENTATION WITH MIXED-SCALE HIERARCHICAL TRANSFORMER
11476	MODELING THE IMPACT OF INTER-RATER DISAGREEMENT ON SLEEP STATISTICS USING DEEP GENERATIVE LEARNING
1778	MODULO SAMPLING AND RECOVERY IN SHIFT-INVARIANT SPACES
4359	MOMA: MIXTURE-OF-MODALITY-ADAPTATIONS FOR TRANSFERRING KNOWLEDGE FROM IMAGE MODELS TOWARDS EFFICIENT AUDIO-VISUAL ACTION RECOGNITION
6785	MOMENTUM-IMBUED LANGEVIN DYNAMICS (MILD) FOR FASTER SAMPLING
11908	MONAI FOR DEEP-LEARNING BASED CBCT RECONSTRUCTION
8096	MONOSTATIC DMG PASSIVE SENSING WITH HYPOTHESIS TESTING
10418	Monte Carlo Self-Training For Speech Recognition
5896	MOS-FAD: IMPROVING FAKE AUDIO DETECTION VIA AUTOMATIC MEAN OPINION SCORE PREDICTION
8844	MOSIC: MULTIMODAL SEMANTIC INTEGRATED COMMUNICATION FOR HEALTH MONITORING IN IOT SCENARIOS
2707	MOSSFORMER2: COMBINING TRANSFORMER AND RNN-FREE RECURRENT NETWORK FOR ENHANCED TIME-DOMAIN MONAURAL SPEECH SEPARATION
7580	Motif-Matching Based Sub-Braingraph Level Networks for Noisy Resting-State fMRI Analysis
6860	MOTION LATENT DIFFUSION FOR STOCHASTIC TRAJECTORY PREDICTION
3524	MOTION TRANSFER-DRIVEN INTRA-CLASS DATA AUGMENTATION FOR FINGER VEIN RECOGNITION
8631	Motion-Tolerant Radar-based Heart Sound Detection
1694	MRTNet: Multi-Resolution Temporal Network for Video Sentence Grounding
9745	MSFD: Multi-Scale Feature Distillation for Semantic Segmentation

<b>Paper Number</b>	<b>Paper Title</b>
7213	MSFR: Stance Detection based on Multi-aspect Semantic Feature Representation via Hierarchical Contrastive Learning
3284	MSG-BART: Multi-granularity Scene Graph-Enhanced Encoder-Decoder Language Model for Video-grounded Dialogue Generation
8488	MS-SENET: ENHANCING SPEECH EMOTION RECOGNITION THROUGH MULTI-SCALE FEATURE FUSION WITH SQUEEZE-AND-EXCITATION BLOCKS
2644	MSSTNET: A MULTI-SCALE SPATIO-TEMPORAL CNN-TRANSFORMER NETWORK FOR DYNAMIC FACIAL EXPRESSION RECOGNITION
11907	MST--: A MODIFICATION OF MST++ FOR NARROW DOMAIN HYPERSPECTRAL RECONSTRUCTION
8182	MTA: A LIGHTWEIGHT MULTILINGUAL TEXT ALIGNMENT MODEL FOR CROSS-LANGUAGE VISUAL WORD SENSE DISAMBIGUATION
3059	MTDIFFUSION: MULTI-TASK DIFFUSION MODEL WITH DUAL-UNET FOR FOLEY SOUND GENERATION
1617	MTIDNET: A MULTIMODAL TEMPORAL INTEREST DETECTION NETWORK FOR VIDEO SUMMARIZATION
4105	MTRGL: Effective Temporal Correlation Discerning through Multi-modal Temporal Relational Graph Learning
2717	MULTI-AGENT 3D SEISMIC EXPLORATION USING ADAPT-THEN-COMBINE FULL WAVEFORM INVERSION IN A HARDWARE-IN-THE-LOOP SYSTEM
1351	MULTI-AGENT EXPLORATION VIA SELF-LEARNING AND SOCIAL LEARNING
3829	MULTI-AGENT SPARSE INTERACTION MODELING IS AN ANOMALY DETECTION PROBLEM
4014	Multi-Antenna ISAC Receiver with n-Tuple Blind Deconvolution
6985	Multi-Attention Enhanced Discriminator for GAN-Based Anomalous Sound Detection
6167	Multi-band speech tensor decomposition for interactive feature extraction in early dysphagia screening

<b>Paper Number</b>	<b>Paper Title</b>
1184	MULTI-BEAM MULTIPLEXING DESIGN WITH PHASE-ONLY EXCITATION BASED ON HYBRID BEAMFORMING ARCHITECTURES
7498	Multicast Transmission Design with Enhanced DoF for MIMO Coded Caching Systems
2169	MULTICAST WITH MULTIPLE WARDENS IN IRS-AIDED COVERT DFRC SYSTEM
2300	MULTI-CHANNEL MOSRA: MEAN OPINION SCORE AND ROOM ACOUSTICS ESTIMATION USING SIMULATED DATA AND A TEACHER MODEL
2097	MULTI-CMGAN+/: LEVERAGING MULTI-OBJECTIVE SPEECH QUALITY METRIC PREDICTION FOR SPEECH ENHANCEMENT
3465	Multi-dimension Queried and Interacting Network for Stereo Image Deraining
9654	MULTI-DIMENSIONAL GEOMETRIC FEATURE-BASED CALIBRATION METHOD FOR LIDAR AND CAMERA FUSION
1551	Multidimensional Scaling-Based TDOA Localization in Modified Polar Representation
4265	Multi-dimensional Speech Quality Assessment in Crowdsourcing
9997	Multi-grained Multimodal Interaction Network for Sentiment Analysis
7624	MULTI-INTEREST LEARNING FOR MULTI-MODAL PAPER RECOMMENDATION
9323	MULTI-LABEL ABNORMALITY CLASSIFICATION FROM 12-LEAD ECG USING A 2D RESIDUAL U-NET
8846	MULTI-LAYER RELATION KNOWLEDGE DISTILLATION FOR FINGERPRINT RESTORATION
6316	MULTI-LEVEL AUGMENTATION CONSISTENCY LEARNING AND SAMPLE SELECTION FOR SEMI-SUPERVISED DOMAIN GENERALIZATION
5449	MULTI-LEVEL CONTRASTIVE LEARNING FOR HYBRID CROSS-MODAL RETRIEVAL
4320	MULTI-LEVEL GRAPH LEARNING FOR AUDIO EVENT CLASSIFICATION AND HUMAN-PERCEIVED ANNOYANCE

<b>Paper Number</b>	<b>Paper Title</b>
	RATING PREDICTION
9371	MULTI-LEVEL SPATIAL-TEMPORAL FEATURE AGGREGATION AND ALIGNMENT-BASED SELECTIVE RESIDUAL DENSE PROPAGATION MODULE FOR HDR VIDEO RECONSTRUCTION
3579	MULTI-LINEAR KERNEL REGRESSION AND IMPUTATION VIA MANIFOLD LEARNING: THE DYNAMIC MRI CASE
7802	MULTILINGUAL AND FULLY NON-AUTOREGRESSIVE ASR WITH LARGE LANGUAGE MODEL FUSION: A COMPREHENSIVE STUDY
2129	MULTILINGUAL AUDIO-VISUAL SPEECH RECOGNITION WITH HYBRID CTC/RNN-T FAST CONFORMER
4061	MULTILINGUAL DISTILWHISPER: EFFICIENT DISTILLATION OF MULTI-TASK SPEECH MODELS VIA LANGUAGE-SPECIFIC EXPERTS
3079	MULTILINGUAL TRANSLITERATION FOR PAN-INDIC KEYBOARD INPUT
2733	MULTI-MICROPHONE NOISE DATA AUGMENTATION FOR DNN-BASED OWN VOICE RECONSTRUCTION FOR HEARABLES IN NOISY ENVIRONMENTS
7398	MULTIMODAL BREATHING RATE ESTIMATION USING FACIAL MOTION AND RPPG FROM RGB CAMERA
4226	MULTI-MODAL CONTINUAL PRE-TRAINING FOR AUDIO ENCODERS
3195	MULTI-MODAL EMOTION RECOGNITION USING MULTIPLE ACOUSTIC FEATURES AND DUAL CROSS-MODAL TRANSFORMER
8684	MULTI-MODAL GPT-4 AIDED ACTION PLANNING AND REASONING FOR SELF-DRIVING VEHICLES
1676	MULTIMODAL GRAPH-BASED AUDIO-VISUAL EVENT LOCALIZATION
7801	MULTIMODAL IMAGING FEATURE EXTRACTION WITH REFERENCE CANONICAL CORRELATION ANALYSIS UNDERLYING INTELLIGENCE
6914	MULTIMODAL MODELING FOR SPOKEN LANGUAGE IDENTIFICATION

<b>Paper Number</b>	<b>Paper Title</b>
5761	Multimodal Multi-view Spectral-Spatial-Temporal Masked Autoencoder for Self-supervised Emotion Recognition
5106	MULTIMODAL SENTIMENT ANALYSIS BASED ON 3D STEREOSCOPIC ATTENTION
10047	MULTIMODAL SURVIVAL ENSEMBLE NETWORK: INTEGRATING GENOMIC AND HISTOPATHOLOGICAL INSIGHTS FOR ENHANCED CANCER PROGNOSIS
10075	Multimodal Transformer Distillation for Audio-Visual Synchronization
4436	MULTIMODAL TRANSFORMER WITH A LOW-COMPUTATIONAL-COST GUARANTEE
5979	MULTI-MODALITY ACTION RECOGNITION BASED ON DUAL FEATURE SHIFT IN VEHICLE CABIN MONITORING
1442	MULTI-MODALITY CONDITIONAL DIFFUSION MODEL FOR TIME SERIES FORECASTING OF LIVE SALES VOLUME
4592	MULTI-MODALITY SPEECH RECOGNITION DRIVEN BY BACKGROUND VISUAL SCENES
8229	MULTI-MODEL WIRELESS FEDERATED LEARNING WITH DOWNLINK BEAMFORMING
5830	MULTI-OBJECT EDITING IN PERSONALIZED TEXT-TO-IMAGE DIFFUSION MODEL VIA SEGMENTATION GUIDANCE
8848	MULTI-OBJECT TRACKING FOR UNMANNED AERIAL VEHICLES BASED ON MULTI-FRAME FEATURE FUSION
8385	MULTI-OBJECTIVE PROGRESSIVE CLUSTERING FOR SEMI-SUPERVISED DOMAIN ADAPTATION IN SPEAKER VERIFICATION
4706	MULTI-PERSON RESPIRATION RATE ESTIMATION WITH SINGLE PAIR OF TRANSMIT AND RECEIVE ANTENNA
4423	MULTIPLE OBJECT TRACKING BASED ON OCCLUSION-AWARE EMBEDDING CONSISTENCY LEARNING
1583	MULTIPLE PLAYER TRACKING WITH 3D PROJECTION AND SPATIO-TEMPORAL INFORMATION IN MULTI-VIEW SPORTS VIDEOS
2072	MULTIPLE REPRESENTATION TRANSFER FROM LARGE LANGUAGE MODELS TO END-TO-END ASR SYSTEMS



<b>Paper Number</b>	<b>Paper Title</b>
9956	MULTI-RATE VARIABLE-LENGTH CSI COMPRESSION FOR FDD MASSIVE MIMO
6268	MULTI-RELATIONAL GRAPH DIFFUSION NEURAL NETWORK WITH PARALLEL RETENTION FOR STOCK TRENDS CLASSIFICATION
5023	Multiscale Attention Distillation for Object Detection
2703	MULTISCALE AUGMENTED NORMALIZING FLOWS FOR IMAGE COMPRESSION
11449	MULTISCALE COARSE-TO-FINE GUIDED SCREENSHOT DEMOIRÉING
3228	MULTI-SCALE FUSION OF GATED NEIGHBORHOOD ATTENTION TRANSFORMERS FOR SINGLE IMAGE DERAISING
7003	MULTISCALE MATCHING DRIVEN BY CROSS-MODAL SIMILARITY CONSISTENCY FOR AUDIO-TEXT RETRIEVAL
9260	MULTI-SCALE PERMUTATION ENTROPY FOR AUDIO DEEPFAKE DETECTION
2332	MULTISCALE SCORING MODEL FOR ENHANCED URBAN PERCEPTION EVALUATION
11495	Multi-Scale Spectral Loss Revisited
3580	MULTI-SCALE SUB-BAND CONSTANT-Q TRANSFORM DISCRIMINATOR FOR HIGH-FIDELITY VOCODER
6829	MULTI-SENSOR MULTI-SCAN RADAR SENSING OF MULTIPLE EXTENDED TARGETS
2513	Multi-Signal Fusion of Social Diffusion Graph with Bi-directional Semantic Consistency
2141	Multi-source DOA estimation with statistical coverage guarantees
2945	Multi-Source Domain Adaptation meets Dataset Distillation through Dataset Dictionary Learning
7893	MULTI-SOURCE DOMAIN ADAPTATION WITH TRANSFORMER-BASED FEATURE GENERATION FOR SUBJECT-INDEPENDENT EEG-BASED EMOTION RECOGNITION
2367	Multi-Source Domain Generalization for ECG-based Cognitive Load Estimation: Adversarial Invariant and

<b>Paper Number</b>	<b>Paper Title</b>
	Plausible Uncertainty Learning
3960	Multi-Source Dynamic Interactive Network Collaborative Reasoning Image Captioning
2176	Multi-Source Unsupervised Transfer Components Learning for Cross-Domain Speech Emotion Recognition
3694	MULTI-SPEAKER LOCALIZATION IN THE CIRCULAR HARMONIC DOMAIN ON SMALL APERTURE MICROPHONE ARRAYS USING DEEP CONVOLUTIONAL NETWORKS
10443	Multispectral Filter Array Design by Optimal Sphere Packing
7927	MULTISPECTRAL RF IMAGING USING MULTIPLE NARROW-BAND FMCW SIGNALS
8668	MULTI-STAGE CONTRASTIVE REGRESSION FOR ACTION QUALITY ASSESSMENT
8721	MULTI-STAGE LEARNING FOR RADAR PULSE ACTIVITY SEGMENTATION
9707	MULTI-STAGE PROGRESSIVE REFINEMENT AND ROI CONTEXT ENHANCEMENT NETWORK FOR SMALL LOGO DETECTION
11878	MULTI-STAGE TRAINING FOR CROSS-DOMAIN FULL-BAND AUDIO PACKET LOSS CONCEALMENT
1486	Multistatic passive detection of cyclostationary signals
11502	Multi-stream Acoustic Modelling using Raw Real and Imaginary Parts of the Fourier Transform
10314	MULTITARGET TRACKING IN THE PRESENCE OF VELOCITY AMBIGUITY FOR AUTOMOTIVE RADAR
10157	MULTI-TASK CASCADED ATTENTION NETWORK FOR BRAIN TUMOR SEGMENTATION AND CLASSIFICATION
4375	MULTITASK CLASSIFICATION OF ANTIMICROBIAL PEPTIDES FOR SIMULTANEOUS ASSESSMENT OF ANTIMICROBIAL PROPERTY AND STRUCTURAL FOLD
4328	MULTI-TASK LEARNING FOR FRONT-END TEXT PROCESSING IN TTS
5284	MULTI-TASK PSEUDO-LABEL LEARNING FOR NON-INTRUSIVE SPEECH QUALITY ASSESSMENT MODEL

<b>Paper Number</b>	<b>Paper Title</b>
3693	MULTI-TASK SELF-SUPERVISED LEARNING FOR MEDICAL IMAGE SEGMENTATION
9419	MULTITASK SPEECH RECOGNITION AND SPEAKER CHANGE DETECTION FOR UNKNOWN NUMBER OF SPEAKERS
2696	MULTI-TEACHER DISTILLATION FOR INCREMENTAL OBJECT DETECTION
8667	MULTIVARIATE DENSITY ESTIMATION USING LOW-RANK FEJER-RIESZ FACTORIZATION
2195	Multivariate Fourier Distribution Perturbation: Domain Shifts with Uncertainty in Frequency Domain
3241	MULTIVARIATE TIME SERIES FORECASTING WITH CAUSAL-TEMPORAL ATTENTION NETWORK
2786	MULTI-VIEW INTERACTIVE COMPROMISE LEARNING FOR GROUP RECOMMENDATION
6478	MULTI-VIEW MIDIVAE: FUSING TRACK- AND BAR-VIEW REPRESENTATIONS FOR LONG MULTI-TRACK SYMBOLIC MUSIC GENERATION
1840	MULTI-VIEW SPEAKER EMBEDDING LEARNING FOR ENHANCED STABILITY AND DISCRIMINABILITY
4979	MULTI-VIEW SPECTROGRAM TRANSFORMER FOR RESPIRATORY SOUND CLASSIFICATION
5130	Multi-view Subspace Clustering with Consensus Graph Contrastive Learning
6471	MULTIWAY-ADAPTER: ADAPTING MULTIMODAL LARGE LANGUAGE MODELS FOR SCALABLE IMAGE-TEXT RETRIEVAL
6789	MULTI-WEATHER DEGRADATION-AWARE TRANSFORMER FOR IMAGE RESTORATION
3437	MUSIC AUTO-TAGGING WITH ROBUST MUSIC REPRESENTATION LEARNED VIA DOMAIN ADVERSARIAL TRAINING
11942	MUSIC ENHANCEMENT WITH DEEP FILTERS: A TECHNICAL REPORT FOR THE ICASSP 2024 CADENZA CHALLENGE
4106	MUSIC SOURCE SEPARATION BASED ON A LIGHTWEIGHT DEEP LEARNING FRAMEWORK (DTTNET: DUAL-PATH TFC-TDF UNET)

<b>Paper Number</b>	<b>Paper Title</b>
3058	MUSIC SOURCE SEPARATION WITH BAND-SPLIT ROPE TRANSFORMER
1482	MUSIC UNDERSTANDING LLAMA: ADVANCING TEXT-TO-MUSIC GENERATION WITH QUESTION ANSWERING AND CAPTIONING
7866	MUSICLDM: ENHANCING NOVELTY IN TEXT-TO-MUSIC GENERATION USING BEAT-SYNCHRONOUS MIXUP STRATEGIES
3920	MUSIC-TO-DANCE POSES: LEARNING TO RETRIEVE DANCE POSES FROM MUSIC
1926	MuSR: Multi-Scale 3D Scenes Reconstruction based on Monocular Video
7235	MUTUAL INFORMATION ASSISTED GRAPH CONVOLUTION NETWORK FOR COLD-START RECOMMENDATION
8722	Mutual information based Noise Scale optimization for Gradient Leakage Resistant Federated Learning
1101	MUTUAL INFORMATION-BASED FAIR ACTIVE LEARNING
1431	MUTUALITY ATTRIBUTE MAKES BETTER VIDEO ANOMALY DETECTION
1544	MUTUALREG: MUTUAL LEARNING FOR UNSUPERVISED MEDICAL IMAGE REGISTRATION
3293	MVITP: MULTI-VIEW IMAGE-TEXT PERCEPTION FOR FEW-SHOT REMOTE SENSING IMAGE CLASSIFICATION
7289	NAC: MITIGATING NOISY CORRESPONDENCE IN CROSS-MODAL MATCHING VIA NEIGHBOR AUXILIARY CORRECTOR
2022	NATURAL LANGUAGE SUPERVISION FOR GENERAL-PURPOSE AUDIO REPRESENTATIONS
7699	NEAR-FIELD LOCALIZATION WITH 1-BIT QUANTIZED HYBRID A/D RECEPTION
9771	NEAR-FIELD MIMO CHANNEL RECONSTRUCTION VIA LIMITED GEOMETRY FEEDBACK
7512	NEAR-FIELD NEURAL RENDERING GUIDED BY SINGLE-SHOT PHOTOMETRIC STEREO
9253	NEBNET:EXPLOITING NODE-EDGE BI-LEVEL NETWORK FOR GENE EXPRESSION PREDICTION

<b>Paper Number</b>	<b>Paper Title</b>
10272	Neighborhood-Enhanced Multimodal Collaborative Filtering for Item Cold Start Recommendation
4699	NeRF-AD: Neural Radiance Field with Attention-based Disentanglement for Talking Face Synthesis
1660	NeRF-Gaze: A Head-Eye Redirection Parametric Model for Gaze Estimation
3977	NERI: IMPLICIT NEURAL REPRESENTATION OF LIDAR POINT CLOUD USING RANGE IMAGE SEQUENCE
4277	NEURAL AMBISONICS ENCODING FOR COMPACT IRREGULAR MICROPHONE ARRAYS
9378	NEURAL CONCATENATIVE SINGING VOICE CONVERSION: RETHINKING CONCATENATION-BASED APPROACH FOR ONE-SHOT SINGING VOICE CONVERSION
1683	Neural Network Training Strategy to Enhance Anomaly Detection Performance: A Perspective on Reconstruction Loss Amplification
9228	NEURAL NETWORK-BASED SYMBOLIC REGRESSION FOR EMPIRICAL MODELING OF THE BEHAVIOR OF A PLANETARY GEARBOX
4784	NEURAL NETWORK-BASED VIRTUAL MICROPHONE ESTIMATION WITH VIRTUAL MICROPHONE AND BEAMFORMER-LEVEL MULTI-TASK LOSS
9761	Neural Ordinary differential equations with Trainable solvers
7056	NEURAL SPEAKER DIARIZATION USING MEMORY-AWARE MULTI-SPEAKER EMBEDDING WITH SEQUENCE-TO-SEQUENCE ARCHITECTURE
7812	Neural Stochastic Differential Equations with Change Points: A Generative Adversarial Approach
8741	NEURAL2SPEECH: A TRANSFER LEARNING FRAMEWORK FOR NEURAL-DRIVEN SPEECH RECONSTRUCTION
6667	NEUROHEED+: IMPROVING NEURO-STEERED SPEAKER EXTRACTION WITH JOINT AUDITORY ATTENTION DETECTION
7762	NEUROMORPHIC SENSING MEETS UNLIMITED SAMPLING
8791	NEW INTENT DISCOVERY WITH MULTI-VIEW CLUSTERING

<b>Paper Number</b>	<b>Paper Title</b>
9186	NEWTONALIZED ORTHOGONAL MATCHING PURSUIT FOR MIXED FAR-FIELD AND NEAR-FIELD SOURCE LOCALIZATION
5282	NEXT-TDNN: MODERNIZING MULTI-SCALE TEMPORAL CONVOLUTION BACKBONE FOR SPEAKER VERIFICATION
7136	NIIRF: NEURAL IIR FILTER FIELD FOR HRTF UPSAMPLING AND PERSONALIZATION
7828	NLSIT: A NON-LOCAL STEREO INTERACTION TRANSFORMER FOR STEREO IMAGE SUPER-RESOLUTION
7573	NOISE MASKING ATTACKS AND DEFENSES FOR PRETRAINED SPEECH MODELS
4046	NOISE2ONE: ONE-SHOT IMAGE DENOISING WITH LOCAL IMPLICIT LEARNING
9122	Noise-Aware Speech Separation with Contrastive Learning
4541	NOISE-BERT: A UNIFIED PERTURBATION-ROBUST FRAMEWORK WITH NOISE ALIGNMENT PRE-TRAINING FOR NOISY SLOT FILLING TASK
3763	NOISE-DISENTANGLED GRAPH CONTRASTIVE LEARNING VIA LOW-RANK AND SPARSE SUBSPACE DECOMPOSITION
9500	NOISE-RESISTANT GRAPH NEURAL NETWORK FOR NODE CLASSIFICATION
7440	NOISE-ROBUST DSP-ASSISTED NEURAL PITCH ESTIMATION WITH VERY LOW COMPLEXITY
6788	NOISE-ROBUST ZERO-SHOT TEXT-TO-SPEECH SYNTHESIS CONDITIONED ON SELF-SUPERVISED SPEECH-REPRESENTATION MODEL WITH ADAPTERS
8038	NOISY IMAGE RESTORATION BASED ON CONDITIONAL ACCELERATION SCORE APPROXIMATION
3333	NOISY-ARCMIX: ADDITIVE NOISY ANGULAR MARGIN LOSS COMBINED WITH MIXUP FOR ANOMALOUS SOUND DETECTION
3029	NoLACE: Improving Low-Complexity Speech Codec Enhancement Through Adaptive Temporal Shaping
7117	NOMAD: UNSUPERVISED LEARNING OF PERCEPTUAL EMBEDDINGS FOR SPEECH ENHANCEMENT AND NON-MATCHING REFERENCE AUDIO QUALITY ASSESSMENT

<b>Paper Number</b>	<b>Paper Title</b>
8623	Non Commutative Convolutional Signal Models in Neural Networks: Stability to Small Deformations
3435	NONASYMPTOTIC PERFORMANCE LIMITS OF LOW-LATENCY SECURE INTEGRATED SENSING AND COMMUNICATION SYSTEMS
1817	NON-INTRUSIVE SPEECH INTELLIGIBILITY PREDICTION FOR HEARING-IMPAIRED USERS USING INTERMEDIATE ASR FEATURES AND HUMAN MEMORY MODELS
5583	NON-INTRUSIVE SPEECH QUALITY ASSESSMENT WITH MULTI-TASK LEARNING BASED ON TENSOR NETWORK
6552	NON-ITERATIVE PYRAMID NETWORK FOR UNSUPERVISED DEFORMABLE MEDICAL IMAGE REGISTRATION
11548	Nonlinear Graph Wavelets via Medianfication
4074	NONLINEARITY DETECTION AND COMPENSATION FOR EEG-BASED SPEECH TRACKING
10189	NON-STATIONARY BANDITS WITH PERIODIC BEHAVIOR: HARNESSING RAMANUJAN PERIODICITY TRANSFORMS TO CONQUER TIME-VARYING CHALLENGES
1042	Non-uniform Frequency Spacing for Regularization-free Gridless DOA
2572	NORMALIZATION IS ALL YOU NEED: ROBUST FULL-RANGE CONTACTLESS SPO2 ESTIMATION ACROSS USERS
2080	NPRF: NEURAL PAINTED RADIOSITY FIELDS FOR NEURAL IMPLICIT RENDERING AND SURFACE RECONSTRUCTION
5770	NTT SPEAKER DIARIZATION SYSTEM FOR CHIME-7: MULTI-DOMAIN, MULTI-MICROPHONE END-TO-END AND VECTOR CLUSTERING DIARIZATION
4964	Nuclear-norm Maximization for Low-Rank Updates
6660	NUV-DOA: NUV PRIOR-BASED BAYESIAN SPARSE RECONSTRUCTION WITH SPATIAL FILTERING FOR SUPER-RESOLUTION DOA ESTIMATION
5088	NWS: NATURAL TEXTUAL BACKDOOR ATTACKS VIA WORD SUBSTITUTION
4451	OADAS: OPTIMIZING GLOBAL PERTURBATION ATTACKS WITH DUAL-PATH ATTRIBUTION SYNERGY

<b>Paper Number</b>	<b>Paper Title</b>
8795	Object Correlation Matrix For Two-Stage Object Detection Network
9700	OBJECT DETECTION ORIENTED PRIVACY-PRESERVING FRAME-LEVEL VIDEO ANOMALY DETECTION
7611	Object Trajectory Estimation with Multi-Band Wi-Fi Neural Dynamic Fusion
1791	OBJECT-CONDITIONED BAG OF INSTANCES FOR FEW-SHOT PERSONALIZED INSTANCE RECOGNITION
5288	ODAQ: OPEN DATASET OF AUDIO QUALITY
8870	OFDM WAVEFORM DESIGN WITH GOOD CORRELATION LEVEL AND PEAK-TO-MEAN ENVELOPE POWER RATIO FOR THE JOINT MIMO RADAR AND COMMUNICATIONS
3341	OFFLINE REINFORCEMENT LEARNING BASED ON NEXT STATE SUPERVISION
1967	OFFLINE REINFORCEMENT LEARNING WITH GENERATIVE ADVERSARIAL NETWORKS AND UNCERTAINTY ESTIMATION
1207	OFFLINE REINFORCEMENT LEARNING WITH POLICY GUIDANCE AND UNCERTAINTY ESTIMATION
5430	OLKAVS: AN OPEN LARGE-SCALE KOREAN AUDIO-VISUAL SPEECH DATASET
5200	OMNIDIRECTIONAL MULTI-ROTOR AERIAL VEHICLE POSE OPTIMIZATION: A NOVEL APPROACH TO PHYSICAL LAYER SECURITY
7277	ON ESTIMATING LINK PREDICTION UNCERTAINTY USING STOCHASTIC CENTERING
5277	ON FINE-TUNING PRE-TRAINED SPEECH MODELS WITH EMA-TARGET SELF-SUPERVISED LOSS
7894	On Generalized Signature Graphs
4929	ON HRTF NOTCH FREQUENCY PREDICTION USING ANTHROPOMETRIC FEATURES AND NEURAL NETWORKS
7182	ON IMPROVED DISTRIBUTED RANDOM RESHUFFLING OVER NETWORKS
10447	ON MEASURES OF UNCERTAINTY IN CLASSIFICATION



<b>Paper Number</b>	<b>Paper Title</b>
10050	ON OPTIMIZING TIMESTEPS OF AN EDM BASED DIFFUSION SAMPLING PROCEDURE
2262	ON REAL-TIME MULTI-STAGE SPEECH ENHANCEMENT SYSTEMS
6973	ON THE CHOICE OF THE OPTIMAL TEMPORAL SUPPORT FOR AUDIO CLASSIFICATION WITH PRE-TRAINED EMBEDDINGS
10449	ON THE CONTRACTIVITY OF PLUG-AND-PLAY OPERATORS
8480	On the Convergence of Hierarchical Federated Learning with Gradient Quantization and Imperfect Transmission
5784	ON THE CONVERGENCE OF SINGLE-TIMESCALE MULTI-SEQUENCE STOCHASTIC APPROXIMATION WITHOUT FIXED POINT SMOOTHNESS
3067	ON THE DESIGN OF PLANAR DIFFERENTIAL MICROPHONE ARRAYS WITH SPECIFIED BEAMWIDTH OR SIDELobe LEVEL
4161	ON THE EFFECT OF DATA-AUGMENTATION ON LOCAL EMBEDDING PROPERTIES IN THE CONTRASTIVE LEARNING OF MUSIC AUDIO REPRESENTATIONS
3450	ON THE EQUIVALENCE OF DYNAMIC MODE DECOMPOSITION AND COMPLEX NONNEGATIVE MATRIX FACTORIZATION
10435	ON THE ESTIMATION OF TSALLIS ENTROPY AND A NOVEL INFORMATION MEASURE BASED ON ITS PROPERTIES
6996	ON THE GENERALIZATION ERROR OF BYZANTINE-RESILIENT DECENTRALIZED LEARNING
8243	ON THE IMPORTANCE OF NEURAL WIENER FILTER FOR RESOURCE EFFICIENT MULTICHANNEL SPEECH ENHANCEMENT
2145	ON THE OPEN PROMPT CHALLENGE IN CONDITIONAL AUDIO GENERATION
8895	ON THE PRIVACY OF FEDERATED CLUSTERING: A CRYPTOGRAPHIC VIEW
9549	ON THE RELATION BETWEEN INTERNAL LANGUAGE MODEL AND SEQUENCE DISCRIMINATIVE TRAINING FOR NEURAL TRANSDUCERS

<b>Paper Number</b>	<b>Paper Title</b>
8808	ON THE RESILIENCE OF ONLINE FEDERATED LEARNING TO MODEL POISONING ATTACKS THROUGH PARTIAL SHARING
6053	ON THE ROLE OF ROOM ACOUSTICS IN AUDIO PRESENTATION ATTACK DETECTION
1782	ON THE TRADEOFF BETWEEN PRIVACY PRESERVATION AND BYZANTINE-ROBUSTNESS IN DECENTRALIZED LEARNING
7825	ON TIME-ENCODED SAMPLING FOR MULTIGENERATOR SHIFT INVARIANT SPACES
11454	On Training Speech Separation Models With Various Numbers of Speakers
9887	ON UNIQUE LOCALIZATION OF UNCORRELATED CONSTANT-MODULUS SOURCES USING SPARSE LINEAR ARRAYS
4649	ON-DEVICE CONSTRAINED SELF-SUPERVISED LEARNING FOR KEYWORD SPOTTING VIA QUANTIZATION AWARE PRE-TRAINING AND FINE-TUNING
7452	ONE MODEL TO RULE THEM ALL ? TOWARDS END-TO-END JOINT SPEAKER DIARIZATION AND SPEECH RECOGNITION
7061	ONE-BIT QUANTIZATION ROBUST TO ANGLE-OF-ARRIVALS FOR UNIFORM LINEAR ANTENNA ARRAY
5599	ONE-CLASS KNOWLEDGE DISTILLATION FOR SPOOFING SPEECH DETECTION
6439	ONE-EPOCH TRAINING WITH SINGLE TEST SAMPLE IN TEST TIME FOR BETTER GENERALIZATION OF COUGH-BASED COVID-19 DETECTION MODEL
5829	ONE-SHOT SENSITIVITY-AWARE MIXED SPARSITY PRUNING FOR LARGE LANGUAGE MODELS
2798	One-stage Deep Stereo Network
2585	ONE-STAGE TRAINING GENERATIVE PARADIGM FOR GENERALIZED ZERO-SHOT LEARNING
10107	One-Step Late Fusion Multi-view Clustering with Compressed Subspace
4250	Online Auditing of Information Flow

<b>Paper Number</b>	<b>Paper Title</b>
5529	Online Caching with Switching Cost and Operational Long-term Constraints: An Online Learning Approach
7964	ONLINE MOUSE BEHAVIOR DETECTION BY HISTORICAL DEPENDENCY AND TYPICAL INSTANCES
6084	ONLINE SPEAKER DIARIZATION OF MEETINGS GUIDED BY SPEECH SEPARATION
3552	ONLINE TARGET SOUND EXTRACTION WITH KNOWLEDGE DISTILLATION FROM PARTIALLY NON-CAUSAL TEACHER
4642	Open-set DeepFake Detection to fight the Unknown
2492	OPENTE: OPEN-STRUCTURE TABLE EXTRACTION FROM TEXT
7122	Open-vocabulary Keyword-spotting with Adaptive Instance Normalization
3426	Open-Vocabulary Skeleton Action Recognition with Diffusion Graph Convolutional Network and Pre-Trained Vision-Language Models
6178	OPINE: Leveraging A Optimization-Inspired Deep Unfolding Method for Multi-channel Speech Enhancement
1081	OPNet: Deep Occlusion Perception Network with Boundary Awareness for Amodal Instance Segmentation
5899	OPTIMAL ANN-SNN CONVERSION WITH GROUP NEURONS
8361	OPTIMAL BEAMFORMING STRUCTURE FOR RATE SPLITTING MULTIPLE ACCESS
3430	OPTIMAL BER MINIMUM PRECODER DESIGN FOR OTFS-BASED ISAC SYSTEMS
8987	OPTIMAL STRUCTURE OF RECEIVE BEAMFORMING FOR OVER-THE-AIR COMPUTATION
9081	Optimal transport distances for directed, weighted graphs: a case study with cell-cell communication networks
3292	OPTIMIZING k IN kNN GRAPHS WITH GRAPH LEARNING PERSPECTIVE
11863	OPTIMIZING MUSIC SOURCE SEPARATION IN COMPLEX AUDIO ENVIRONMENTS THROUGH PROGRESSIVE SELF-KNOWLEDGE DISTILLATION

<b>Paper Number</b>	<b>Paper Title</b>
7554	Optimizing Synchronization Delay for Digital Twin over Wireless Networks
7325	Optimizing Trading Strategies in Quantitative Markets using Multi-Agent Reinforcement Learning
11473	OSIN: OBJECT-CENTRIC SCENE INFERENCE NETWORK FOR UNSUPERVISED VIDEO ANOMALY DETECTION
11464	Outlier Censoring via Block Sparse Learning
5508	OUTLIER-ROBUST FEATURE SELECTION WITH L2,1-NORM MINIMIZATION AND GROUP ROW-SPARSITY INDUCED CONSTRAINTS
4104	OUT-OF-DISTRIBUTION DETECTION FOR LEARNING-BASED CHEST X-RAY DIAGNOSIS
8514	P2DT: MITIGATING FORGETTING IN TASK-INCREMENTAL LEARNING WITH PROGRESSIVE PROMPT DECISION TRANSFORMER
3131	PaCaS-WAA: Patch-based Contrastive Semi-supervised Learning with Wavelet Guidance and Adaptive Augmentation for Tumour Segmentation
9352	PANORAMIC IMAGE INPAINTING WITH GATED CONVOLUTION AND CONTEXTUAL RECONSTRUCTION LOSS
2522	PARALINGUISTICS-ENHANCED LARGE LANGUAGE MODELING OF SPOKEN DIALOGUE
5527	PARALLEL AUGMENTATION AND DUAL ENHANCEMENT FOR OCCLUDED PERSON RE-IDENTIFICATION
7737	PARAMETER EFFICIENT AUDIO CAPTIONING WITH FAITHFUL GUIDANCE USING AUDIO-TEXT SHARED LATENT REPRESENTATION
4724	Parameter Efficient Finetuning for Speech Emotion Recognition and Domain Adaptation
11540	PARAMETER ESTIMATION PROCEDURES FOR DEEP MULTI-FRAME MVDR FILTERING FOR SINGLE-MICROPHONE SPEECH ENHANCEMENT
4272	Parameter Estimation via Expectation Maximization - Expectation Consistent Algorithm

<b>Paper Number</b>	<b>Paper Title</b>
4263	PARAMETER-EFFICIENT ADAPTATION FOR COMPUTATIONAL IMAGING
6716	Pareto Graph Self-Supervised Learning
7632	PARODY DETECTION USING SOURCE-TARGET ATTENTION WITH TEACHER-FOURCED LYRICS
1407	PART REPRESENTATION LEARNING WITH TEACHER-STUDENT DECODER FOR OCCLUDED PERSON RE-IDENTIFICATION
9555	PARTIAL CONVOLUTIONAL BASED-RADIO MAP RECONSTRUCTION FOR URBAN ENVIRONMENTS WITH INACCESSIBLE AREAS
3848	PARTIALLY OBSERVABLE MODEL-BASED LEARNING FOR ISAC RESOURCE ALLOCATION
7180	PASTE AND HARMONIZE VIA DENOISING: SUBJECT-DRIVEN IMAGE EDITING WITH FROZEN PRE-TRAINED DIFFUSION MODEL
7080	Patch Inherent Feature Guided Mask Selection for Image Compression
9039	PATCH-LEVEL KNOWLEDGE DISTILLATION AND REGULARIZATION FOR MISSING MODALITY MEDICAL IMAGE SEGMENTATION
2425	Patch-wise Augmentation for Anomaly Detection and Localization
7892	PATIENT-ADAPTIVE AND LEARNED MRI DATA UNDERSAMPLING USING NEIGHBORHOOD CLUSTERING
11882	PATIENT-SPECIFIC MODELING OF DAILY ACTIVITY PATTERNS FOR UNSUPERVISED DETECTION OF PSYCHOTIC AND NON-PSYCHOTIC RELAPSES
9772	PAVITS: EXPLORING PROSODY-AWARE VITS FOR END-TO-END EMOTIONAL VOICE CONVERSION
3602	PECER: EMPATHETIC RESPONSE GENERATION VIA DYNAMIC PERSONALITY EXTRACTION AND CONTEXTUAL EMOTIONAL REASONING
7142	PECR: PARAMETER-EFFICIENT TRANSFER LEARNING WITH CROSS-MODAL REPRESENTATION LEARNING FOR REMOTE SENSING VISUAL QUESTION ANSWERING

<b>Paper Number</b>	<b>Paper Title</b>
10442	PENDANTSS: PEnalized Norm-Ratios Disentangling Additive Noise, Trend and Sparse Spikes
7727	PERCEIVING MULTI-LAYER REPRESENTATIONS FOR NO-REFERENCE IMAGE QUALITY ASSESSMENT
9460	PERCEPTUAL QUALITY EVALUATION FOR FASTER PLAYBACK VIDEOS
7563	PERCEPTUALLY-MOTIVATED SPATIAL AUDIO CODEC FOR HIGHER-ORDER AMBISONICS COMPRESSION
7476	PERFORMANCE AND ENERGY BALANCE: A COMPREHENSIVE STUDY OF STATE-OF-THE-ART SOUND EVENT DETECTION SYSTEMS
1322	PERFORMANCE CONDITIONING FOR DIFFUSION-BASED MULTI-INSTRUMENT MUSIC SYNTHESIS
4832	PERIOULAR BIOMETRICS ENHANCEMENT THROUGH MULTIMODAL EMBEDDINGS AND CLASSIFIER ADAPTATION
10206	PERIODGRAD: TOWARDS PITCH-CONTROLLABLE NEURAL VOCODER BASED ON A DIFFUSION PROBABILISTIC MODEL
3813	Permutation-alignment method using manifold optimization for frequency-domain blind source separation
5955	PERSONA EXTRACTION THROUGH SEMANTIC SIMILARITY FOR EMOTIONAL SUPPORT CONVERSATION GENERATION
11926	PERSONALISED ANOMALY DETECTORS AND PROTOTYPICAL REPRESENTATIONS FOR RELAPSE DETECTION FROM WEARABLE-BASED DIGITAL PHENOTYPING
1861	PERSONALIZATION OF CTC-BASED END-TO-END SPEECH RECOGNITION USING PRONUNCIATION-DRIVEN SUBWORD TOKENIZATION
7654	Personalized Federated Learning with Attention-based Client Selection
6534	PERSONALIZED LOCAL DIFFERENTIALLY PRIVATE FEDERATED LEARNING WITH ADAPTIVE CLIENT SAMPLING
7026	PERSONALIZED NEURAL SPEECH CODEC

<b>Paper Number</b>	<b>Paper Title</b>
7254	PERSONALIZED OVER-THE-AIR FEDERATED LEARNING WITH PERSONALIZED RECONFIGURABLE INTELLIGENT SURFACES
9692	PFCF-NET: A NETWORK BASED ON PROGRESSIVE FEATURE INTERACTION AND CROSS-SCALE FEATURE FUSION FOR REMOTE SENSING CHANGE DETECTION
6961	PFDM: Parser-Free Virtual Try-on via Diffusion Model
6854	PHASE CONTINUITY-AWARE SELF-ATTENTIVE RECURRENT NETWORK WITH ADAPTIVE FEATURE SELECTION FOR ROBUST VAD
8473	PHASE LEARNING BASED ON INTERACTIVE PERCEPTION FOR LIMITED-SAMPLE RESIDENTIAL AREA SEMANTIC SEGMENTATION
9853	PHASE RECONSTRUCTION IN SINGLE CHANNEL SPEECH ENHANCEMENT BASED ON PHASE GRADIENTS AND ESTIMATED CLEAN-SPEECH AMPLITUDES
5621	Phase Retrieval by Tensor Total Least Squares
8349	PHASE-SPACE-GUIDED DEEP LEARNING FOR TIME SERIES FORECASTING
7528	PHISANET: PHONETICALLY INFORMED SPEECH ANIMATION NETWORK
3643	PHONEME-AWARE ENCODING FOR PREFIX-TREE-BASED CONTEXTUAL ASR
9357	Photovoltaic power forecasting using sky images and sun motion
7986	PhyOT: Physics-informed object tracking in surveillance cameras
7354	Physically-constrained block-term tensor decomposition for polarimetric image recovery
10448	PHYSICS-GUIDED DEEP SCATTER ESTIMATION BY WEAK SUPERVISION FOR QUANTITATIVE SPECT
7718	PHYSICS-GUIDED VARIATIONAL GRAPH AUTOENCODER FOR AIR QUALITY INFERENCE
8103	PIANO TRANSCRIPTION WITH HARMONIC ATTENTION
9195	PILOT LENGTH MINIMIZATION VIA AP-UE CLUSTERING IN CELL-FREE SYSTEMS

<b>Paper Number</b>	<b>Paper Title</b>
7242	PIXEL-SUPERPIXEL CONTRASTIVE LEARNING AND PSEUDO-LABEL CORRECTION FOR HYPERSPECTRAL IMAGE CLUSTERING
4255	PJSCC: A PUNCTURING-BASED JOINT SOURCE CHANNEL CODING SCHEME WITH HIERARCHICAL DOWN-SAMPLING LAYER
5337	PLS: UNSUPERVISED DOMAIN ADAPTATION FOR 3D OBJECT DETECTION VIA PSEUDO-LABEL SIZES
6350	Plug-and-Play Algorithm coupled with Low-Rank Quadratic Envelope Regularization for Compressive Spectral Imaging
8926	PLUG-AND-PLAY MVDR BEAMFORMING FOR SPEECH SEPARATION
8203	PMDI: COMBINING PARAMETRIC-MODEL AND DEPTH-AWARE IMPLICIT FUNCTION FOR SINGLE-VIEW HUMAN RECONSTRUCTION
1129	PMMWDECONV: UNSUPERVISED DATA-CONSISTENT BLIND PASSIVE MILLIMETER-WAVE IMAGE DECONVOLUTION WITH GLOBAL CONTEXT PRIORS
4655	PN-DetX: A Dedicated Framework for Pulmonary Nodule Detection in X-ray Images
6972	POISONING-FREE DEFENSE AGAINST BLACK-BOX MODEL EXTRACTION
10042	PoisonPrompt: Backdoor Attack on Prompt-based Large Language Models
2520	POLARDB: FORMULA-DRIVEN DATASET FOR PRE-TRAINING TRAJECTORY ENCODERS
6506	POLITICAL TWEET SENTIMENT ANALYSIS FOR PUBLIC OPINION POLLING
11451	PoP-IDLMA: Product-of-Prior Independent Deeply Learned Matrix Analysis for Multichannel Music Source Separation
1920	POSE-HMR: HEURISTIC TRANSFORMER WITH POSTURAL PRIOR CONSTRAINTS FOR 3D HUMAN MESH RECONSTRUCTION
7411	POSITION-AWARE ACTIVE LEARNING FOR MULTI-MODAL ENTITY ALIGNMENT



<b>Paper Number</b>	<b>Paper Title</b>
9264	POSITIVE TRANSFER OF THE WHISPER SPEECH TRANSFORMER TO HUMAN AND ANIMAL VOICE ACTIVITY DETECTION
8990	Posterior Sampling Algorithms for Unsupervised Speech Enhancement with Recurrent Variational Autoencoder
7153	POSTERIOR VARIANCE-PARAMETERISED GAUSSIAN DROPOUT: IMPROVING DISENTANGLED SEQUENTIAL AUTOENCODERS FOR ZERO-SHOT VOICE CONVERSION
4387	POST-TRAINING EMBEDDING ALIGNMENT FOR DECOUPLING ENROLLMENT AND RUNTIME SPEAKER RECOGNITION MODELS
3608	POWER-AWARE TASK-BASED LEARNING OF NEUROMORPHIC ADCS
8933	PRACTICAL CHALLENGE AND SOLUTION FOR IRS-AIDED INDOOR LOCALIZATION SYSTEM
1188	Predict and Interpret Health Risk using EHR through Typical Patients
1252	PREDICTING ADVERSE EVENTS FOR PATIENTS WITH TYPE-1 DIABETES VIA SELF-SUPERVISED LEARNING
3275	PREDICTING FALL EVENTS BY A SPATIO-TEMPORAL TOPOLOGICAL NETWORK WITH MULTIPLE WEARABLE SENSORS
9462	PREDICTING RTMS TREATMENT EFFECTS USING OPEN-LOOP CONTROL AND NEURAL MANIFOLD
4809	PREDICTION-CORRECTION LINE SEGMENT DETECTION
3501	PRE-ECHO REDUCTION IN TRANSFORM AUDIO CODING VIA TEMPORAL ENVELOPE CONTROL WITH MACHINE LEARNING BASED ESTIMATION
3205	PRE-POST INTERACTION LEARNING FOR BRAIN TUMOR SEGMENTATION WITH MISSING MRI MODALITIES
6656	PRE-TRAINED ACOUSTIC-AND-TEXTUAL MODELING FOR END-TO-END SPEECH-TO-TEXT TRANSLATION
7891	PRIORITIZING DATA ACQUISITION FOR END-TO-END SPEECH MODEL IMPROVEMENT
2330	PRIVACY LEAKAGE IN GRAPH SIGNAL TO GRAPH MATCHING PROBLEMS

<b>Paper Number</b>	<b>Paper Title</b>
7767	PRIVACY PRESERVING FEDERATED LEARNING FROM MULTI-INPUT FUNCTIONAL PROXY RE-ENCRYPTION
4709	PRIVACY PRESERVING GAZE ESTIMATION VIA FEDERATED LEARNING ADAPTED TO EGOCENTRIC VIDEO
5131	PRIVACY-AWARE JOINT SOURCE-CHANNEL CODING FOR IMAGE TRANSMISSION BASED ON DISENTANGLED INFORMATION BOTTLENECK
9114	PRIVACY-PRESERVING ATTENTION-WEIGHTED MULTI-SOURCE DOMAIN ADAPTATION FOR EEG MOTOR IMAGERY
3997	PRIVACY-PRESERVING DEEP LEARNING USING DEFORMABLE OPERATORS FOR SECURE TASK LEARNING
3365	PRIVACY-PRESERVING DISTRIBUTED OPTIMISATION USING STOCHASTIC PDMM
2790	ProAug: Prototype-Based Augmentation for Long-Tailed Image Classification
7250	PROBABILISTIC SIMPLEX COMPONENT ANALYSIS VIA VARIATIONAL AUTO-ENCODING
3084	PROBABILISTIC SPIKE TRAIN INFERENCE
9538	PROBABILITY-AWARE WORD-CONFUSION-NETWORK-TO-TEXT ALIGNMENT APPROACH FOR INTENT CLASSIFICATION
1465	PROBMCL: SIMPLE PROBABILISTIC CONTRASTIVE LEARNING FOR MULTI-LABEL VISUAL CLASSIFICATION
7602	PROFILE-ERROR-TOLERANT TARGET-SPEAKER VOICE ACTIVITY DETECTION
9230	PROGRESSIVE IMAGE SYNTHESIS FROM SEMANTICS TO DETAILS WITH DENOISING DIFFUSION GAN
4637	PROGRESSIVE LEARNING BASED KNOWLEDGE DISTILLATION FOR LOW RESOLUTION CEREBRAL MICROBLEED SEGMENTATION
6762	PROGRESSIVE UNSUPERVISED DOMAIN ADAPTATION FOR ASR USING ENSEMBLE MODELS AND MULTI-STAGE TRAINING
10076	PROGRESSIVELY LEARNING FROM MACRO-EXPRESSIONS FOR MICRO-EXPRESSION RECOGNITION

<b>Paper Number</b>	<b>Paper Title</b>
4128	PRO-HAN: A HETEROGENEOUS GRAPH ATTENTION NETWORK FOR PROFILE-BASED SPOKEN LANGUAGE UNDERSTANDING
2103	Promoting Independence of Depression and Speaker Features for Speaker Disentanglement in Speech-based Depression Detection
3308	PROMPTASR FOR CONTEXTUALIZED ASR WITH CONTROLLABLE STYLE
4110	PROMPT-BASED PERSONALIZED FEDERATED LEARNING FOR MEDICAL VISUAL QUESTION ANSWERING
4917	Prompt-driven Target Speech Diarization
7403	PROMPTFORMER: PROMPTED CONFORMER TRANSDUCER FOR ASR
7690	PROMPTING AUDIOS USING ACOUSTIC PROPERTIES FOR EMOTION REPRESENTATION
6562	PROMPTING LABEL EFFICIENCY IN FEDERATED GRAPH LEARNING VIA PERSONALIZED SEMI-SUPERVISION
5572	PROMPTING LARGE LANGUAGE MODELS WITH FINE-GRAINED VISUAL RELATIONS FROM SCENE GRAPH FOR VISUAL QUESTION ANSWERING
7953	Prompting Large Language Models with Speech Recognition Abilities
2407	PROMPTING TO PROMPT FOR REHEARSAL-FREE CLASS INCREMENTAL LEARNING
9691	PROMPTTTS++: CONTROLLING SPEAKER IDENTITY IN PROMPT-BASED TEXT-TO-SPEECH USING NATURAL LANGUAGE DESCRIPTIONS
3424	PROMPTVC: FLEXIBLE STYLISTIC VOICE CONVERSION IN LATENT SPACE DRIVEN BY NATURAL LANGUAGE PROMPTS
7395	PROPOSAL DISTILLATION OF MULTI-MODAL FEATURE AGGREGATION NETWORK FOR VIDEO OBJECT DETECTION
4990	PROTOTYPE CALIBRATION WITH SYNTHESIZED SAMPLES FOR ZERO-SHOT CHINESE CHARACTER RECOGNITION
4239	Prototype-Guided Masking for Unsupervised Domain Adaptation

<b>Paper Number</b>	<b>Paper Title</b>
3083	PROVABLE RANDOMIZED COORDINATE DESCENT FOR MATRIX COMPLETION
3612	PROXIMAL BELLMAN MAPPINGS FOR REINFORCEMENT LEARNING AND THEIR APPLICATION TO ROBUST ADAPTIVE FILTERING
1445	PseKD: Phase-shift Encoded Knowledge Distillation for Oriented Object Detection in Remote Sensing Images
5049	Pseudo Labels Regularization for Imbalanced Partial-Label Learning
3112	Pseudo-outlier synthesis using q-Gaussian distributions for out-of-distribution detection
4770	PU-EdgeFormer++: an Advanced Hierarchical Edge Transformer for Arbitrary-Scale Point Cloud Upsampling using Distance Fields
6619	PUSH4REC: TEMPORAL AND CONTEXTUAL TREND-AWARE TRANSFORMER PUSH NOTIFICATION RECOMMENDER
4806	PVCG: Prompt-based Vision-aware Classification and Generation for Multi-modal Rumor Detection
3883	PVITNET: AN EFFECTIVE APPROACH FOR ANDROID MALWARE DETECTION USING PYRAMID FEATURE PROCESSING AND VISION TRANSFORMER
4752	Pyramid: A Heterogeneous Data Integration Algorithm Based on Hierarchical Graph
4403	QUANTIFYING SPATIAL AUDIO QUALITY IMPAIRMENT
4329	QUANTIFYING THE EFFECT OF SIMULATOR-BASED DATA AUGMENTATION FOR SPEECH RECOGNITION ON AUGMENTED REALITY GLASSES
8040	QUANTIZATION NOISE MASKING IN PERCEPTUAL NEURAL AUDIO CODER
7922	QUANTIZED DECODER IN LEARNED IMAGE COMPRESSION FOR DETERMINISTIC RECONSTRUCTION
11472	QUANTIZED RADIO MAP ESTIMATION USING TENSOR AND DEEP GENERATIVE MODELS
11556	Quantum Algorithm for Signal Denoising
8261	QUANTUM FEDERATED LEARNING WITH QUANTUM NETWORKS

<b>Paper Number</b>	<b>Paper Title</b>
6031	QUANTUM INSPIRED IMAGE AUGMENTATION APPLICABLE TO WAVEGUIDES AND OPTICAL IMAGE TRANSFER VIA ANDERSON LOCALIZATION
7487	QUANTUM PRIVACY AGGREGATION OF TEACHER ENSEMBLES (QPATE) FOR PRIVACY PRESERVING QUANTUM MACHINE LEARNING
5632	Quantum Ranging Enhanced TDoA Localization
3834	QUANTUM TOPIC MODEL: TOPIC MODELING USING VARIATIONAL QUANTUM CIRCUITS
8059	QUAPPROX: A FRAMEWORK FOR BENCHMARKING THE APPROXIMABILITY OF VARIATIONAL QUANTUM CIRCUIT
7612	RADAR PERCEPTION WITH SCALABLE CONNECTIVE TEMPORAL RELATIONS FOR AUTONOMOUS DRIVING
4550	RADAR RECOGNITION IN THE WILD: ENHANCING RADAR EMITTER RECOGNITION THROUGH AUTO-CORRELATION MODEL-AGNOSTIC META LEARNING
10288	RadarDiff: Improving Sea Clutter Suppression using Diffusion Models for Radar images
6006	RADEMACHER COMPLEXITY REGULARIZATION FOR CORRELATION-BASED MULTIVIEW REPRESENTATION LEARNING
7065	RADIO SLAM WITH HYBRID SENSING FOR MIXED REFLECTION TYPE ENVIRONMENTS
11886	RAD-NET: A REPAIRING AND DENOISING NETWORK FOR SPEECH SIGNAL IMPROVEMENT
2121	RANDOMIZED MAXIMUM LIKELIHOOD VIA HIGH-DIMENSIONAL BAYESIAN OPTIMIZATION
9442	RANKING ENHANCED FINE-GRAINED CONTRASTIVE LEARNING FOR RECOMMENDATION
4772	RANKING OF VISUAL TRACKERS USING ROBUST ERROR NORMS
10132	RAPID CHANGE LOCALIZATION IN DYNAMIC GRAPHICAL MODELS
1279	RAPID HYBRID MODULAR RECEIVE BEAMFORMING VIA LEARNED OPTIMIZATION

<b>Paper Number</b>	<b>Paper Title</b>
8940	Rate-Quality based Rate Control Model for Neural Video Compression
9663	RATING-AUGMENTED NO-REFERENCE POINT CLOUD QUALITY ASSESSMENT USING MULTI-TASK LEARNING
2244	RCIF: TOWARDS ROBUST DISTRIBUTED DNN COLLABORATIVE INFERENCE UNDER HIGHLY LOSSY NETWORKS
4703	RDANET:REJECT DOMAIN ATTENTION NETWORK FOR CONFUSED FACIAL EXPRESSION RECOGNITION
4933	RD-COST REGRESSION SPEED UP TECHNIQUE FOR VVC INTRA BLOCK PARTITIONING
4677	RD-NeRF: Neural Robust Distilled Feature Fields for Sparse-view Scene Segmentation
1571	READ, SPELL AND REPEAT: SCENE TEXT RECOGNITION WITH VISION-LANGUAGE CIRCULAR REFINEMENT
9860	REAL-ORIENTED OBJECT DETECTION DRIVEN BY INTELLIGENT STOCKBREEDING
3927	REAL-TIME LOW-LATENCY MUSIC SOURCE SEPARATION USING HYBRID SPECTROGRAM-TASNET
1598	REAL-TIME MULTI-HUMAN PARSING ON EMBEDDED DEVICES
4684	REAL-TIME PRIVACY-PRESERVING FALL RISK ASSESSMENT WITH A SINGLE BODY-WORN TRACKING CAMERA
4364	REAL-TIME STEREO SPEECH ENHANCEMENT WITH SPATIAL-CUE PRESERVATION BASED ON DUAL-PATH STRUCTURE
11901	REBUILD, REGENERATE: A GATED TEMPORAL CONVOLUTION BASED GAN FOR SPEECH SIGNAL IMPROVEMENT
7661	RECAP: RETRIEVAL-AUGMENTED AUDIO CAPTIONING
7577	RECENT ADVANCES IN SCALABLE ENERGY-EFFICIENT AND TRUSTWORTHY SPIKING NEURAL NETWORKS: FROM ALGORITHMS TO TECHNOLOGY
2342	RECOGNITION-GUIDED DIFFUSION MODEL FOR SCENE TEXT IMAGE SUPER-RESOLUTION
10052	Reconstruction of sound field through diffusion models

<b>Paper Number</b>	<b>Paper Title</b>
4267	Recovering from Privacy-Preserving Masking with Large Language Models
10187	RECOVERING MISSING NODE FEATURES WITH LOCAL STRUCTURE-BASED EMBEDDINGS
7749	RECURSIVE-TAIL-FISTA FOR SPARSE SIGNAL RECOVERY
3005	Redefining Night Vision: The Power of MSR-Driven Neural ISP
2876	REDUCED-DIMENSIONAL DECOMPOSITION AND EIGENSPACE RECONSTRUCTION OF COHERENT SOURCES WITH ARBITRARY RECTANGLE ARRAYS
6438	REDUCING THE COMPLEXITY OF NORMALIZING FLOW ARCHITECTURES FOR POINT CLOUD ATTRIBUTE COMPRESSION
3961	REFERENCE LINE NETWORK: ON SIMULTANEOUS GAUSSIAN LINE DETECTION AND CONNECTION GRAPH INFERENCE
3863	Refinement Bird's Eye View Feature for 3D Lane Detection with Dual-Branch View Transformation Module
4460	REFINING 3D HUMAN MESH VIA MODEL-FREE OFFSETS ESTIMATION
7829	Refining Text Input for Augmentative and Alternative Communication (AAC) Devices: Analysing Language Model Layers for Optimisation
9298	REFLECTION REMOVAL USING RECURRENT POLARIZATION-TO-POLARIZATION NETWORK
3207	REFLOW-TTS: A RECTIFIED FLOW MODEL FOR HIGH-FIDELITY TEXT-TO-SPEECH
2869	REGION-ADAPTIVE VIDEO SHARPENING VIA RATE-PERCEPTION OPTIMIZATION
2287	REGIR: REFINED GEOMETRY FOR SINGLE-IMAGE IMPLICIT CLOTHED HUMAN RECONSTRUCTION
3016	REGULARIZED CONDITIONAL ALIGNMENT FOR MULTI-DOMAIN TEXT CLASSIFICATION
9545	Regularized Contrastive Pre-training for Few-shot Bioacoustic Sound Detection

<b>Paper Number</b>	<b>Paper Title</b>
5249	Reinforcement Learning Compensated Filter for Multi-agents Cooperative Localization
7631	REINFORCEMENT LEARNING-GUIDED OPTOGENETIC STIMULATION POLICIES FOR ROBUST FUNCTIONAL NETWORK DISCOVERY
9602	RELATIONAL GRAPH-BRIDGED IMAGE-TEXT INTERACTION: A NOVEL METHOD FOR MULTI-MODAL RELATION EXTRACTION
4840	REMIXED2REMIXED: DOMAIN ADAPTATION FOR SPEECH ENHANCEMENT BY NOISE2NOISE LEARNING WITH REMIXING
11930	REMIXING MUSIC FOR HEARING AIDS USING ENSEMBLE OF FINE-TUNED SOURCE SEPARATORS
11879	RENET: A TIME-FREQUENCY DOMAIN GENERAL SPEECH RESTORATION NETWORK FOR ICASSP 2024 SPEECH SIGNAL IMPROVEMENT CHALLENGE
3570	RENYI DIFFERENTIAL PRIVACY IN THE SHUFFLE MODEL: ENHANCED AMPLIFICATION BOUNDS
9064	RENYI DIVERGENCES LEARNING FOR EXPLAINABLE CLASSIFICATION OF SAR IMAGE PAIRS
4680	REPARAMETERIZATION HEAD FOR EFFICIENT MULTI-INPUT NETWORKS
3950	Representation and Boundary enhancement for Action Segmentation using Transformer
1433	REPRESENTATION LEARNING ACROSS FEATURE AND TOPOLOGY VIEWS WITH OUTPUT CORRECTION FOR GRAPH CONVOLUTIONAL NETWORKS
11560	Representation Learning With Hidden Unit Clustering For Low Resource Speech Applications
4308	REPURPOSING MU-MIMO DOWNLINK FOR JOINT WIRELESS COMMUNICATIONS AND IMAGING VIA VIRTUAL USERS
9451	RESIDUAL DENSE SWIN TRANSFORMER FOR CONTINUOUS DEPTH-INDEPENDENT ULTRASOUND IMAGING
5123	RESIDUALTRANSFORMER: RESIDUAL LOW-RANK LEARNING WITH WEIGHT-SHARING FOR TRANSFORMER



<b>Paper Number</b>	<b>Paper Title</b>
	LAYERS
2812	RESOURCE-CONSTRAINED STEREO SINGING VOICE CANCELLATION
4462	RESOURCE-EFFICIENT SEPARATION TRANSFORMER
2994	RETAINING INFORMATIVE LATENT VARIABLES IN PROBABILISTIC SEGMENTATION
5310	RETHINKING NORMALS: DIRECTION GUIDED POINT CLOUD RECOGNITION
8616	RETHINKING SESSION VARIABILITY: LEVERAGING SESSION EMBEDDINGS FOR SESSION ROBUSTNESS IN SPEAKER VERIFICATION
8260	RETHINKING TARGETED ADVERSARIAL ATTACKS FOR NEURAL MACHINE TRANSLATION
7942	Retrieval Augmented End-to-End Spoken Dialog Models
3735	RETRIEVAL-AUGMENTED TEXT-TO-AUDIO GENERATION
7125	RETRIEVAL-GENERATION SYNERGY AUGMENTED LARGE LANGUAGE MODELS
7971	REVEALING EMOTIONAL CLUSTERS IN SPEAKER EMBEDDINGS: A CONTRASTIVE LEARNING STRATEGY FOR SPEECH EMOTION RECOGNITION
5314	REVERSIBLE JUMP MARKOV CHAIN MONTE CARLO FOR PULSE FITTING
4662	REVISE THE NLU: A PROMPTING STRATEGY FOR ROBUST DIALOGUE SYSTEM
11557	Revisiting Deep Generalized Canonical Correlation Analysis
7934	REVISITING SELF-SUPERVISED LEARNING OF SPEECH REPRESENTATION FROM A MUTUAL INFORMATION PERSPECTIVE
8966	REVISITING THE EQUIVALENCE OF IN-CONTEXT LEARNING AND GRADIENT DESCENT: THE IMPACT OF DATA DISTRIBUTION
10286	REWEIGHTED ATOMIC NORM MINIMIZATION FOR ONE-BIT MULTICHANNEL SPECTRAL COMPRESSED SENSING
2444	RGB IMAGES ENHANCING HYPERSPECTRAL IMAGE DENOISING WITH DIFFUSION MODEL

<b>Paper Number</b>	<b>Paper Title</b>
11912	RGBT2HS-Net: Reconstructing a hyper-spectral volume from an RGB-T stack via an attention-powered multiresolution framework
7813	Riemannian Diffusion Adaptation over Graphs with Application to Online Distributed PCA
4021	RIS LOCALIZATION AND SPATIALLY WIDEBAND FILTERING EFFECTS
8454	RISK-MANAGED SPARSE INDEX TRACKING VIA MARKET GRAPH CLUSTERING
3267	RK-core: An established methodology for exploring the hierarchical structure within datasets
2273	RL-EMO: A REINFORCEMENT LEARNING FRAMEWORK FOR MULTIMODAL EMOTION RECOGNITION
1900	RL-LOGO: Deep Reinforcement Learning Localization for Logo Recognition
9483	ROBUST AND IMPERCEPTIBLE COMMERCIAL CAMERA-SCREEN COMMUNICATION WITH 60HZ REFRESH RATE
1519	ROBUST BEAMFORMING FOR DFRC SYSTEMS IN COMPLEX ENVIRONMENTS
7352	ROBUST CROSS-DOMAIN SPEAKER VERIFICATION WITH MULTI-LEVEL DOMAIN ADAPTERS
9769	Robust decoding of the auditory attention from EEG recordings through graph convolutional networks
8702	Robust DOA estimation from deep acoustic imaging
3440	ROBUST FACE RECOGNITION BASED ON AN ANGLE-AWARE LOSS AND MASKED AUTOENCODER PRE-TRAINING
8792	Robust Lightweight Depth Estimation Model via Data-free Distillation
10333	ROBUST LOCALIZATION OF KEY FOB USING CHANNEL IMPULSE RESPONSE OF ULTRA WIDE BAND SENSORS FOR KEYLESS ENTRY SYSTEM
3052	ROBUST LOW-RANK CORRELATION FITTING
11459	ROBUST MULTISTATIC TARGET LOCALIZATION IN THE PRESENCE OF NLOS ERRORS AND OUTLIERS

<b>Paper Number</b>	<b>Paper Title</b>
7671	ROBUST NEAR-FIELD BEAMFORMING FOR MILLIMETER WAVE COMMUNICATION SYSTEM WITH APERTURE PERTURBATION
9566	Robust Recovery of Joint Sparse signals via Simultaneous Orthogonal Matching Pursuit
4288	Robust regression analysis based on the K-divergence
1804	ROBUST SELF-SUPERVISED LEARNING WITH CONTRAST SAMPLES FOR NATURAL LANGUAGE UNDERSTANDING
2577	ROBUST SINGLE-PARTICLE CRYO-EM IMAGE DENOISING AND RESTORATION
6210	ROBUST SPEAKER PERSONALISATION USING GENERALIZED LOW-RANK ADAPTATION FOR AUTOMATIC SPEECH RECOGNITION
2036	ROBUST SPOOF SPEECH DETECTION BASED ON MULTI-SCALE FEATURE AGGREGATION AND DYNAMIC CONVOLUTION
9479	ROBUST SYMBOL-LEVEL PRECODING VIA A SYMBOL-PERTURBED ZERO-FORCING STRUCTURE
6965	ROBUST WAKE WORD SPOTTING WITH FRAME-LEVEL CROSS-MODAL ATTENTION BASED AUDIO-VISUAL CONFORMER
3074	ROBUSTNESS AGAINST ADVERSARIAL ATTACKS VIA LEARNING CONFINED ADVERSARIAL POLYTOPES
4746	ROBUSTNESS EVALUATION OF MACHINE LEARNING MODELS FOR ROBOT ARM ACTION RECOGNITION IN NOISY ENVIRONMENTS
8013	ROBUSTTSVAR: A ROBUST TIME SERIES VARIANCE ESTIMATION ALGORITHM
8106	RoFi: Robust WiFi Intrusion Detection via Distribution Matching
11552	ROTOR NOISE-AWARE NOISE COVARIANCE MATRIX ESTIMATION FOR UNMANNED AERIAL VEHICLE AUDITION
5602	RSED: Zero-shot Relation Triplet Extraction via Relation Selection and Entity Boundary Detection
4886	RTLBP - AN EFFICIENT LOCAL PATTERN FOR FACIAL IMAGES RETRIEVAL

<b>Paper Number</b>	<b>Paper Title</b>
11482	RTSNet: Learning to Smooth in Partially Known State-Space Models
3179	RVAE-EM: GENERATIVE SPEECH DEREVERBERATION BASED ON RECURRENT VARIATIONAL AUTO-ENCODER AND CONVOLUTIVE TRANSFER FUNCTION
3936	RVDNET: A TWO-STAGE NETWORK FOR REAL-WORLD VIDEO DESNOWING WITH DOMAIN ADAPTATION
3048	S2E: Towards an End-to-End Entity Resolution Solution from Acoustic Signal
2432	SADA: SAUDI AUDIO DATASET FOR ARABIC
9848	SADE: A Speaker-Aware Dual Encoding Model based on DiagBERT for Medical Triage and Pre-diagnosis
6598	SALIENCY PREDICTION OF SPORTS VIDEOS: A LARGE-SCALE DATABASE AND A SELF-ADAPTIVE APPROACH
10374	SALM: Speech-augmented Language Model with In-context Learning for Speech Recognition and Translation
10140	SAM: A Self-adaptive Attention Module for Context-Aware Recommendation System
3926	SAM-Deblur: Let Segment Anything Boost Image Deblurring
7131	SAMF: SMALL-AREA-AWARE MULTI-FOCUS IMAGE FUSION FOR OBJECT DETECTION
9231	SAM-GEBD : ZERO-COST APPROACH FOR GENERIC EVENT BOUNDARY DETECTION
3359	SAM-GUIDED ENHANCED FINE-GRAINED ENCODING WITH MIXED SEMANTIC LEARNING FOR MEDICAL IMAGE CAPTIONING
3893	SAM-OCTA: A Fine-Tuning Strategy for Applying Foundation Model to OCTA Image Segmentation Tasks
6906	Sampling and Recovery of Signals over Product Cell Structures
9879	SAMVG: A MULTI-STAGE IMAGE VECTORIZATION MODEL WITH THE SEGMENT-ANYTHING MODEL
3109	Sandwiched Lo-res Simulation for Scalable Flood Modeling
7206	SAR2NDVI: Pre-training for SAR-to-NDVI Image Translation

<b>Paper Number</b>	<b>Paper Title</b>
1122	SASA: Saliency-Aware Self-Adaptive Snapshot Compressive Imaging
1320	SA-SOT: Speaker-Aware Serialized Output Training for Multi-Talker ASR
1276	SBM: Smoothness-based Minimization for Domain Generalization
7914	SCALABLE AND EFFICIENT SPEECH ENHANCEMENT USING MODIFIED COLD DIFFUSION: A RESIDUAL LEARNING APPROACH
4850	Scalable Ensemble-based Detection Method Against Adversarial Attacks For Speaker Verification
3234	Scalable Model-Based Gaussian Process Clustering
3298	SCALE-AWARE COMPETITION NETWORK FOR PALMPRINT RECOGNITION
2553	Scale-free and Task-generic Attack: Generating Photo-realistic Adversarial Patterns with Patch Quilting Generator
11939	SCALING NVIDIA's MULTI-SPEAKER MULTI-LINGUAL TTS SYSTEMS WITH ZERO-SHOT TTS TO INDIC LANGUAGES
9656	Scaling Results for Robust Distributed Estimation in Sensor Networks using Order Statistics
9481	SCANPCGC: LEARNING-BASED LOSSLESS POINT CLOUD GEOMETRY COMPRESSION USING SEQUENTIAL SLICE REPRESENTATION
6876	SCENE SKETCH-TO-IMAGE SYNTHESIS BASED ON MULTI-OBJECT CONTROL
8692	SC-MAD: MIXTURES OF HIGHER-ORDER NETWORKS FOR DATA AUGMENTATION
8187	SCNet: Sparse Compression Network for Music Source Separation
8784	SCORE CALIBRATION BASED ON CONSISTENCY MEASURE FACTOR FOR SPEAKER VERIFICATION
7973	SCORE: SELF-SUPERVISED CORRESPONDENCE FINE-TUNING FOR IMPROVED CONTENT REPRESENTATIONS
4582	SCORE-BASED DIFFUSION MODELS FOR PHOTOACOUSTIC TOMOGRAPHY IMAGE RECONSTRUCTION

<b>Paper Number</b>	<b>Paper Title</b>
2138	ScoreDec: A Phase-preserving High-Fidelity Audio Codec with A Generalized Score-based Diffusion Post-filter
9216	SCRN: A SPECTROGRAM CONVOLUTIONAL RECURRENT NETWORK FOR AOA ESTIMATION USING BLUETOOTH 5
3615	SDEMG: Score-Based Diffusion Model for Surface Electromyographic Signal Denoising
7954	SD-HUBERT: SENTENCE-LEVEL SELF-DISTILLATION INDUCES SYLLABIC ORGANIZATION IN HUBERT
2120	SDIF-DA: A SHALLOW-TO-DEEP INTERACTION FRAMEWORK WITH DATA AUGMENTATION FOR MULTI-MODAL INTENT DETECTION
8641	SDRNET: SALIENCY-GUIDED DYNAMIC RESTORATION NETWORK FOR RAIN AND HAZE REMOVAL IN NIGHTTIME IMAGES
2679	SEACO-PARAFORMER: A NON-AUTOREGRESSIVE ASR SYSTEM WITH FLEXIBLE AND EFFECTIVE HOTWORD CUSTOMIZATION ABILITY
8122	SEA-GNN: SEQUENCE EXTENSION AUGMENTED GRAPH NEURAL NETWORK FOR SEQUENTIAL RECOMMENDATION
3253	SEAM MASK GUIDED PARTIAL RECONSTRUCTION WITH QUANTUM-INSPIRED LOCAL AGGREGATION FOR DEEP IMAGE STITCHING
4059	Search for gravitational wave probes - A self-supervised learning for pulsars based on signal contexts
1671	Search Robust and Adaptable Architecture
8008	SEC2SEC CO-ATTENTION TRANSFORMER FOR VIDEO-BASED APPARENT AFFECTIVE PREDICTION
7771	SECP: A SPEECH ENHANCEMENT-BASED CURATION PIPELINE FOR SCALABLE ACQUISITION OF CLEAN SPEECH
4101	SECTOR-BASED INTERFERENCE CANCELLATION FOR ROBUST KEYWORD SPOTTING APPLICATIONS USING AN INFORMED MPDR BEAMFORMER
6264	SECURE ENERGY EFFICIENCY FAIRNESS MAXIMIZATION IN BACKSCATTER THROUGHPUT CONSTRAINED UAV-ASSISTED DATA COLLECTION

<b>Paper Number</b>	<b>Paper Title</b>
2859	SECURELY AND EFFICIENTLY OUTSOURCING NEURAL NETWORK INFERENCE VIA PARALLEL MSB EXTRACTION
3935	Security Equivalence Assessment Between Cloud Standards by Mapping of Control Items
9572	SEEING THROUGH THE CONVERSATION: AUDIO-VISUAL SPEECH SEPARATION BASED ON DIFFUSION MODEL
1060	SEEKING SIMILARITIES WHILE REMOVING DIFFERENCES: GRAPH NEURAL NETWORKS BASED ON NODE CORRELATION
8529	SEF-VC: Speaker Embedding Free Zero-Shot Voice Conversion with Cross Attention
6120	SEGLLM: TOPIC-ORIENTED CALL SEGMENTATION VIA LLM-BASED CONVERSATION SYNTHESIS
3513	Segment Anything Model guided Semantic Knowledge Learning for Remote Sensing Change Detection
1277	Segment Anything Model Meets Image Harmonization
5603	SEGMENT THEN MATCH: FIND THE CARRIER BEFORE REASONING IN SCENE-TEXT VQA
9054	SEGMENTATION-DRIVEN INFRARED AND VISIBLE IMAGE FUSION VIA TRANSFORMER-ENHANCED ARCHITECTURE SEARCHING
1844	Segmented Error Minimisation (SEMI) for Robust Training of Deep Learning Models with Non-linear Shifts in Reference Data
9774	SELECTING N-LOWEST SCORES FOR TRAINING MOS PREDICTION MODELS
11549	Selective Acoustic Feature Enhancement for Speech Emotion Recognition with Noisy Speech
10128	Selective Domain-invariant Feature for Generalizable Deepfake Detection
10113	SELECTIVE USER FORWARDED CELL-FREE MASSIVE MIMO WITH QUANTIZED SYMBOLS
8871	SELF KNOWLEDGE DISTILLATION BASED ON LAYER-WISE WEIGHTED FEATURE IMITATION FOR EFFICIENT OBJECT DETECTION

<b>Paper Number</b>	<b>Paper Title</b>
9196	SELF-ADAPTIVE SCALE HANDLING FOR FORECASTING TIME SERIES WITH SCALE HETEROGENEITY
3704	SELF-DISTILLED DYNAMIC FUSION NETWORK FOR LANGUAGE-BASED FASHION RETRIEVAL
1748	SELF-KNOWLEDGE DISTILLATION WITH LEARNING FROM ROLE-MODEL SAMPLES
10421	SELF-MOTION AS SUPERVISION FOR EGOCENTRIC AUDIOVISUAL LOCALIZATION
10238	SELF-SUPERVISED ADAPTIVE AV FUSION MODULE FOR PRE-TRAINED ASR MODELS
6576	SELF-SUPERVISED ADAPTIVE PRE-TRAINING OF MULTILINGUAL SPEECH MODELS FOR LANGUAGE AND DIALECT IDENTIFICATION
3908	Self-supervised Cross-level Consistency Learning for Fundus Image Classification
3117	Self-supervised Domain Exploration with an Optimal Transport Regularization for Open Set Cross-domain Speech Emotion Recognition
8394	SELF-SUPERVISED DUAL GENERATIVE NETWORKS FOR EDGE-PRESERVING IMAGE SMOOTHING
2323	Self-Supervised Face Image Restoration with a One-Shot Reference
1174	SELF-SUPERVISED LEARNING FOR ANOMALOUS SOUND DETECTION
3847	SELF-SUPERVISED LEARNING FOR SLEEP STAGE CLASSIFICATION WITH TEMPORAL AUGMENTATION AND FALSE NEGATIVE SUPPRESSION
7943	SELF-SUPERVISED MODELS OF SPEECH INFER UNIVERSAL ARTICULATORY KINEMATICS
2320	SELF-SUPERVISED MULTI-SCALE HIERARCHICAL REFINEMENT METHOD FOR JOINT LEARNING OF OPTICAL FLOW AND DEPTH
6725	Self-Supervised Path Planning in UAV-aided Wireless Networks based on Active Inference
1954	SELF-SUPERVISED PRETRAINING FOR ROBUST PERSONALIZED VOICE ACTIVITY DETECTION IN ADVERSE



<b>Paper Number</b>	<b>Paper Title</b>
	CONDITIONS
8422	SELF-SUPERVISED PULSE-AWARE INTERPRETABLE DISENTANGLED ECG REPRESENTATION LEARNING
9990	SELF-SUPERVISED REINFORCEMENT LEARNING FOR OUT-OF-DISTRIBUTION RECOVERY VIA AUXILIARY REWARD
8787	SELF-SUPERVISED SPATIALLY VARIANT PSF ESTIMATION FOR ABERRATION-AWARE DEPTH-FROM-DEFOCUS
9492	Self-supervised Speaker Verification Employing a Novel Clustering Algorithm
8148	SELF-SUPERVISED SPEAKER VERIFICATION WITH ADAPTIVE THRESHOLD AND HIERARCHICAL TRAINING
11937	SELF-SUPERVISED SPEECH REPRESENTATION AND CONTEXTUAL TEXT EMBEDDING FOR MATCH-MISMATCH CLASSIFICATION WITH EEG RECORDING
3806	SELF-TRAINING DOMAIN ADAPTATION VIA WEIGHT TRANSMISSION BETWEEN GENERATORS
6979	SELM: Speech Enhancement Using Discrete Tokens and Language Models
6618	SEMANTIC DISTILLATION AND STRUCTURAL ALIGNMENT NETWORK FOR FAKE NEWS DETECTION
7037	SEMANTIC ENRICHMENT FOR VIDEO QUESTION ANSWERING WITH GATED GRAPH NEURAL NETWORKS
8812	SEMANTIC LATENT DECOMPOSITION WITH NORMALIZING FLOWS FOR FACE EDITING
3505	SEMANTIC PROXIMITY ALIGNMENT: TOWARDS HUMAN PERCEPTION-CONSISTENT AUDIO TAGGING BY ALIGNING WITH LABEL TEXT DESCRIPTION
8640	SEMANTIC RECONSTRUCTION OF CONTINUOUS LANGUAGE FROM MEG SIGNALS
4067	SEMANTIC SECURITY: A DIGITAL WATERMARK METHOD FOR IMAGE SEMANTIC PRESERVATION
8878	SEMANTIC SEGMENTATION FOR MULTI-SCENE REMOTE SENSING IMAGES WITH NOISY LABELS BASED ON UNCERTAINTY PERCEPTION
4126	SEMANTIC-ENHANCED SUPERVISED CONTRASTIVE LEARNING

<b>Paper Number</b>	<b>Paper Title</b>
7108	SEMANTIC-GUIDED NETWORK WITH CONTRASTIVE LEARNING FOR VIDEO CAPTION
7738	SEMANTICMAPPER: REGION-SPECIFIC DOMAIN ADAPTATION FOR 3D SHAPES THROUGH LEXICAL DELINEATION
9406	SEMANTIC-PRESERVING IMAGE CODING BASED ON CONDITIONAL DIFFUSION MODELS
7389	SEMANTICS DRIVEN MULTI-VIEW KNOWLEDGE GRAPH EMBEDDING FOR CROSS-LINGUAL ENTITY ALIGNMENT
8712	SemDA: Communication-efficient Data Aggregation Through Distributed Semantic Transmission
7155	SEMI-AUTOREGRESSIVE STREAMING ASR WITH LABEL CONTEXT
1626	SEMI-BLIND ESTIMATION OF DIRECT-TO-REVERBERANT ENERGY RATIO USING RESIDUAL ENERGY TEST STATISTICS
1169	SEMI-DECOUPLED 6D POSE ESTIMATION VIA MULTI-MODAL FEATURE FUSION
3897	SEMI-SUPERVISED DOMAIN ADAPTATION FOR EEG-BASED SLEEP STAGE CLASSIFICATION
5606	SEMI-SUPERVISED METRICS-BASED SELF-TRAINING ROOT CAUSE ANALYSIS FOR CLOUD-NATIVE SYSTEMS WITH CLASS-IMBALANCED DATA
1168	SEMI-SUPERVISED SOUND EVENT DETECTION WITH LOCAL AND GLOBAL CONSISTENCY REGULARIZATION
5960	SEMI-SUPERVISED VOLUMETRIC MEDICAL IMAGE SEGMENTATION VIA CLASS PROTOTYPE GUIDED DISTRIBUTION-ALIGNED REPRESENTATION LEARNING
1768	SENSI-BERT: TOWARDS SENSITIVITY DRIVEN FINE-TUNING FOR PARAMETER-EFFICIENT LANGUAGE MODEL
3355	SENSING WITH RANDOM SIGNALS
4525	SENSING-AIDED COMMUNICATION CHANNEL ESTIMATION WITH TENSOR-BASED MOVING TARGET LOCALIZATION
4791	SENSING-ASSISTED DISTRIBUTED USER SCHEDULING AND BEAMFORMING IN MULI-CELL MMWAVE NETWORKS
8331	SEQUENCE OF LINEAR PROGRAM FOR ROBUST PHASE RETRIEVAL

<b>Paper Number</b>	<b>Paper Title</b>
1855	SEQUENTIAL ACQUISITION OF FEATURES AND EXPERTS FOR DATUM-WISE CLASSIFICATION
8374	SEQUENTIAL DETECTION OF ANOMALIES IN NOISY OUTPUTS OF AN UNKNOWN FUNCTION USING GAUSSIAN AND YULE-SIMON PROCESSES
9431	SEQUENTIAL MONTE CARLO GRAPH CONVOLUTIONAL NETWORK FOR DYNAMIC BRAIN CONNECTIVITY
4108	Sequential Wasserstein Uncertainty Sets for Minimax Robust Online Change Detection
4377	SERC-GCN: SPEECH EMOTION RECOGNITION IN CONVERSATION USING GRAPH CONVOLUTIONAL NETWORKS
7106	SE-SIS: shadow-embeddable lossless secret image sharing for greyscale images
1883	S-Evaluator: Enhance Factual Consistency Evaluator with Adversarial Data Synthesized by Large Language Model
9029	SG2SC: A GENERATIVE SEMANTIC COMMUNICATION FRAMEWORK FOR SCENE UNDERSTANDING-ORIENTED IMAGE TRANSMISSION
7011	SGM: A DATASET FOR 3D GARMENT RECONSTRUCTION FROM SINGLE HAND-DRAWN SKETCH
3865	SGT: SELF-GUIDED TRANSFORMER FOR FEW-SHOT SEMANTIC SEGMENTATION
4558	Shapley Value Guided Extractive Text Summarization
1609	SHIFT OPERATOR AND SEPARATION FILTER FOR DIFFERENT PERIOD MIXED SIGNALS USING COMPANION MATRIX
8806	Shifted-rectangle-window Based Transformer for Non-displaced Femoral Neck Fracture Diagnosis
3520	SIANet: Support Information-Aware Network for Category-Agnostic Pose Estimation
3212	SICRN: ADVANCING SPEECH ENHANCEMENT THROUGH STATE SPACE MODEL AND INPLACE CONVOLUTION TECHNIQUES
5064	SiFiSinger: A High-Fidelity End-to-End Singing Voice Synthesizer Based on Source-Filter Model

<b>Paper Number</b>	<b>Paper Title</b>
7227	SIGNAL RECONSTRUCTION FROM NONIDEAL SAMPLES IN FRACTIONAL FOURIER TRANSFORM DOMAIN
11924	SIGNAL SEPARATION IN RADIO SPECTRUM USING SELF-ATTENTION MECHANISM
2510	Signal Transformer: Complex-valued Attention and Meta-Learning for Signal Recognition
7031	SIGNIFICANT ASR ERROR DETECTION FOR CONVERSATIONAL VOICE ASSISTANTS
6399	SimFall: A Data Generator For RF-based Fall Detection
4181	SIMILAR BUT FASTER: MANIPULATION OF TEMPO IN MUSIC AUDIO EMBEDDINGS FOR TEMPO PREDICTION AND SEARCH
3324	SIMILARITY KNOWLEDGE DISTILLATION WITH CALIBRATED MASK
2868	SimMKD: Simple Mask-Flow Keypoint Detection for both typhoon detection and typhoon eye location
4039	Simple Contrastive Representation Learning for Time Series Forecasting
2054	SIMULTANEOUS INTERIOR AND EXTERIOR SOUND FIELD SYNTHESIS USING CYLINDRICAL AND SPHERICAL LOUDSPEAKER ARRAYS
2507	SIMULTANEOUS POSITIONING AND TRACKING USING DYNAMIC FACTOR GRAPHS AND GEOMETRIC AVERAGE FUSION
8175	SINGFAKE: SINGING VOICE DEEPFAKE DETECTION
3964	SINGLE AND FEW-STEP DIFFUSION FOR GENERATIVE SPEECH ENHANCEMENT
10444	SINGLE CHANNEL MULTIPLE SIGNAL CLASSIFICATION USING PSEUDO-DOPPLER
2995	Single Image Reflection Removal using Feature Difference Enhancement
5941	SINGLE-CHANNEL BLIND DEREVERBERATION BASED ON RANK-1 MATRIX LIFTING IN TIME-FREQUENCY DOMAIN
7773	Single-pixel imaging of dynamic flows using Neural ODE regularization

<b>Paper Number</b>	<b>Paper Title</b>
2647	SINGLE-SOURCE DOMAIN GENERALIZATION IN FUNDUS IMAGE SEGMENTATION VIA MODERATING AND INTERPOLATING INPUT SPACE AUGMENTATION
11944	SINGLE-STAGE TTS WITH ADAPTED VOCODER AND CROSS-ATTENTION: TALTECH SYSTEMS FOR THE LIMMITS'24 CHALLENGE
11862	SIR-PROGRESSIVE AUDIO-VISUAL TF-GRIDNET WITH ASR-AWARE SELECTOR FOR TARGET SPEAKER EXTRACTION IN MISP 2023 CHALLENGE
8113	SITUATIONAL SIGNAL PROCESSING WITH ECOLOGICAL MOMENTARY ASSESSMENT: LEVERAGING ENVIRONMENTAL CONTEXT FOR COCHLEAR IMPLANT USERS
4870	Situation-aware adaptive transmit beamforming for automotive radars
1663	SJTU-TMQA: A QUALITY ASSESSMENT DATABASE FOR STATIC MESH WITH TEXTURE MAP
2622	SKETCH-BASED 3D SHAPE RETRIEVAL WITH MULTI-VIEW FUSION TRANSFORMER
8606	SKETCHED COLUMN-BASED MATRIX APPROXIMATION WITH SIDE INFORMATION
7002	SKILLNET-X: A MULTILINGUAL MULTITASK MODEL WITH SPARSELY ACTIVATED SKILLS
1630	SKIN TONE DISENTANGLEMENT IN 2D MAKEUP TRANSFER WITH GRAPH NEURAL NETWORKS
8075	SKIP-STEP CONTRASTIVE PREDICTIVE CODING FOR TIME SERIES ANOMALY DETECTION
4908	SLIDESPEECH: A LARGE SCALE SLIDE-ENRICHED AUDIO-VISUAL CORPUS
7609	SLOWFAST NETWORK FOR CONTINUOUS SIGN LANGUAGE RECOGNITION
9464	SMALL OBJECT DETECTION ON THE WATER SURFACE BASED ON RADAR AND CAMERA FUSION
1684	SMALL-FOOTPRINT AUTOMATIC SPEECH RECOGNITION SYSTEM USING TWO-STAGE TRANSFER LEARNING BASED SYMMETRIZED TERNARY WEIGHT NETWORK

<b>Paper Number</b>	<b>Paper Title</b>
8447	Small-Footprint Convolutional Neural Network with reduced feature map for Voice Activity Detection
10412	SMMA-NET: AN AUDIO CLUE-BASED TARGET SPEAKER EXTRACTION NETWORK WITH SPECTROGRAM MATCHING AND MUTUAL ATTENTION
3490	SMOOTH START: A UNIFIED APPROACH FOR GRADUAL TRANSITION FROM COLD TO OLD IN RECOMMENDER SYSTEMS
7711	SNAPSHOT PROMPT ENSEMBLE FOR PARAMETER-EFFICIENT SOFT PROMPT TRANSFER
6044	SNORE SOUND FEATURES BASED ON PERCUSSIVE ENHANCING AND POSITIONAL ENCODING COMBINED WITH MULTI-TASK LEARNING FOR OSAHS DETECTION
1951	SOCIAL LEARNING WITH ADAPTIVE MODELS
2289	SOCIAL LODE: HUMAN TRAJECTORY PREDICTION WITH LATENT ODES
9670	SOD-UAV: small object detection for unmanned aerial vehicle images via improved YOLOv7
4812	SOFT ALIGNMENT OF MODALITY SPACE FOR END-TO-END SPEECH TRANSLATION
2112	SOFT DYNAMIC TIME WARPING WITH VARIABLE STEP WEIGHTS
4441	Soft Image Segmentation using Gradient Graph Laplacian Regularizer
10088	SOLUTION AND ANALYSIS FOR 3-D LOCALIZATION IN CLOSED-FORM INTEGRATING SA AND TDOA MEASUREMENTS
6686	SO-NET: MODEL-AGNOSTIC SEQUENTIAL HAND POSE OPTIMIZATION FRAMEWORK
3651	SORTING, REASONING, AND EXTRACTION: AN EASY-TO-HARD REASONING FRAMEWORK FOR DOCUMENT-LEVEL EVENT ARGUMENT EXTRACTION
11505	SOUND FIELD INTERPOLATION FOR ROTATION-INVARIANT MULTICHANNEL ARRAY SIGNAL PROCESSING
1099	SOUNDLOC: AN EFFICIENT CONDITIONAL DISCRETE CONTRASTIVE LATENT DIFFUSION MODEL FOR TEXT-TO-

<b>Paper Number</b>	<b>Paper Title</b>
	SOUND GENERATION
8177	SOURCE-FREE DOMAIN ADAPTATION FOR MILLIMETER WAVE RADAR BASED HUMAN ACTIVITY RECOGNITION
8547	SOURCE-FREE ONLINE DOMAIN ADAPTIVE SEMANTIC SEGMENTATION OF SATELLITE IMAGES UNDER IMAGE DEGRADATION
1959	SourceP: Detecting Ponzi Schemes on Ethereum with Source Code
4289	Space-Time Adaptive Processing for radars in Connected and Automated Vehicular Platoons
8467	SPARSE BAYESIAN LEARNING-BASED DIRECT LOCALIZATION FOR DISTRIBUTED SENSOR ARRAYS WITH UNKNOWN GAIN AND PHASE ERRORS
6319	SPARSE BAYESIAN SYNTHETIC APERTURE PROCESSING BASED DOA ESTIMATION WITH DEFORMED TOWED ARRAYS
1310	SPARSE CHANNEL REPRESENTATION AND ESTIMATION IN NEAR FIELD COMMUNICATIONS
7685	SPARSE PCA WITH FALSE DISCOVERY RATE CONTROLLED VARIABLE SELECTION
4522	Sparse Regularization based on Reverse Ordered Weighted L1-norm and Its Application to Edge-preserving Smoothing
8856	SPARSE SOUND FIELD REPRESENTATION USING COMPLEX ORTHOGONAL MATCHING PURSUIT
2353	SPARSE, WEIGHT-CONSTRAINED ARRAYS WITH $O(N)$ APERTURE FOR REDUCED MUTUAL COUPLING
7276	SPARSELY SHARED LORA ON WHISPER FOR CHILD SPEECH RECOGNITION
5620	SparseSpikformer: A Co-Design Framework for Token and Weight Pruning in Spiking Transformer
1740	SPASE: SPATial Saliency Explanation for time series models
9346	SPATIAL FORMATION-GUIDED NETWORK FOR GROUP ACTIVITY RECOGNITION
7940	SPATIAL SCAPER: A LIBRARY TO SIMULATE AND AUGMENT SOUNDSCAPES FOR SOUND EVENT LOCALIZATION AND DETECTION IN REALISTIC ROOMS

<b>Paper Number</b>	<b>Paper Title</b>
7567	SPATIALCODEC: NEURAL SPATIAL SPEECH CODING
2493	SPATIAL-TEMPORAL INTERACTION DECODING TRANSFORMER FOR UNSUPERVISED MULTIVARIATE TIME SERIES ANOMALY DETECTION
5845	SPATIO-TEMPORAL ACTION DETECTION WITH A MOTION SENSE AND SEMANTIC CORRECTION FRAMEWORK
4629	SPATIO-TEMPORAL CORRELATION LEARNING FOR MULTIPLE OBJECT TRACKING
1750	SPATIO-TEMPORAL DATA MINING WITH INFORMATION INTEGRITY PROTECTION: GRAPH SIGNAL BASED AIR QUALITY PREDICTION
8002	SPATIOTEMPORAL GROUP ANOMALY DETECTION VIA GRAPH TOTAL VARIATION ON TENSORS
3144	SPCL-MER: SUPERVISED PROTOTYPICAL CONTRASTIVE LEARNING FOR MICRO-EXPRESSION RECOGNITION
8831	SPDG-NET: SEMANTICS PRESERVING DOMAIN AUGMENTATION THROUGH STYLE INTERPOLATION FOR MULTI-SOURCE DOMAIN GENERALIZATION
7688	SPEAK WHILE YOU THINK: STREAMING SPEECH SYNTHESIS DURING TEXT GENERATION
3090	SPEAKER ADAPTATION FOR ENHANCEMENT OF BONE-CONDUCTED SPEECH
6612	SPEAKER ANONYMIZATION USING NEURAL AUDIO CODEC LANGUAGE MODELS
11474	SPEAKER ANONYMIZATION USING ORTHOGONAL HOUSEHOLDER NEURAL NETWORK
2942	SPEAKER-ADAPTIVE LIPREADING VIA SPATIO-TEMPORAL INFORMATION LEARNING
10193	SPEAKER-CENTRIC MULTIMODAL FUSION NETWORKS FOR EMOTION RECOGNITION IN CONVERSATIONS
7017	SPECDIFF-GAN: A SPECTRALLY-SHAPED NOISE DIFFUSION GAN FOR SPEECH AND MUSIC SYNTHESIS
4942	SPEC-NERF: MULTI-SPECTRAL NEURAL RADIANCE FIELDS
10030	SPECTRAL ANALYSIS OF VOWELS AND FRICATIVES AT VARIED LEVELS OF DYSARTHRIA SEVERITY FOR AMYOTROPHIC LATERAL SCLEROSIS



<b>Paper Number</b>	<b>Paper Title</b>
10077	Spectral Graph Neural Networks with Generalized Laguerre Approximation
8764	SPECTROGRAM SMOOTHING FOR ESTIMATION OF THE EVOLUTIONARY SPECTRA OF UNIFORMLY MODULATED PROCESSES
10158	SPECTRO-SPATIAL HYPERSPECTRAL IMAGE RECONSTRUCTION FROM INTERFEROMETRIC ACQUISITIONS
8098	SPECTRUMNET: SPECTRUM-BASED TRAJECTORY ENCODE NEURAL NETWORK FOR PEDESTRIAN TRAJECTORY PREDICTION
7800	SPEECH COLLAGE: CODE-SWITCHED AUDIO GENERATION BY COLLAGING MONOLINGUAL CORPORA
11559	Speech Dereverberation With Frequency Domain Autoregressive Modeling
7753	Speech Emotion Recognition with Distilled Prosodic and Linguistic Affect Representations
7064	Speech enhancement in hearing aids using target speech presence estimation based on a delayed remote microphone signal
9293	SPEECH FOUNDATION MODELS ON INTELLIGIBILITY PREDICTION FOR HEARING-IMPAIRED LISTENERS
9194	SPEECH GUIDED MASKED IMAGE MODELING FOR VISUALLY GROUNDED SPEECH
3726	SPEECH RELATIONSHIP LEARNING FOR CROSS-CORPUS SPEECH EMOTION RECOGNITION
7074	Speech Swin-Transformer: Exploring a Hierarchical Transformer with Shifted Windows for Speech Emotion Recognition
9016	SPEECHDPR: END-TO-END SPOKEN PASSAGE RETRIEVAL FOR OPEN-DOMAIN SPOKEN QUESTION ANSWERING
1088	SPEECH-DRIVEN EMOTIONAL 3D TALKING FACE ANIMATION USING EMOTIONAL EMBEDDINGS
2604	SPGFUSION: A SEMANTIC PRIOR GUIDED INFRARED AND VISIBLE IMAGE FUSION NETWORK

<b>Paper Number</b>	<b>Paper Title</b>
1940	SPGM: Prioritizing local features for enhanced speech separation performance
4476	Spiking Structured State Space Model for Monaural Speech Enhancement
9045	Spiking-LEAF: A Learnable Auditory front-end for Spiking Neural Networks
9532	SPIRAL SHAPE MATTERS: NOVEL BIO-INSPIRED COCHLEAR CEPSTRUM
8186	SponTTS: modeling and transferring spontaneous style for TTS
9243	SPOOFING ATTACK AUGMENTATION: CAN DIFFERENTLY-TRAINED ATTACK MODELS IMPROVE GENERALISATION?
6542	SPTESLEEPNET: AUTOMATIC SLEEP STAGING MODEL BASED ON STRIP PATCH EMBEDDINGS AND TRANSFORMER ENCODER
1491	SPY-WATERMARK: ROBUST INVISIBLE WATERMARKING FOR BACKDOOR ATTACK
2917	SRCodec: Split-residual vector quantization for neural speech codec
9489	SRECT: Machine-specific Spatial-resolution Enhancement in Computed Tomography
7012	SR-HUBERT : AN EFFICIENT PRE-TRAINED MODEL FOR SPEAKER VERIFICATION
1520	SRP-UOD: MULTI-BRANCH HYBRID NETWORK FRAMEWORK BASED ON STRUCTURAL RE-PARAMETERIZATION FOR UNDERWATER SMALL OBJECT DETECTION
8774	SR-VFA: ACCURATE SELF-REFINED FACE ALIGNMENT IN VIDEOS
1336	SSHNN: SEMI-SUPERVISED HYBRID NAS NETWORK FOR ECHOCARDIOGRAPHIC IMAGE SEGMENTATION
6294	SSL-NET: A SYNERGISTIC SPECTRAL AND LEARNING-BASED NETWORK FOR EFFICIENT BIRD SOUND CLASSIFICATION
1770	SSR-GPCST: DEEP LEARNING MODELS BASED ON FUNCTIONAL CONNECTIVITY MAPS IN AUTISM RESEARCH
3842	SSTA: Salient Spatially Transformed Attack

<b>Paper Number</b>	<b>Paper Title</b>
7454	STABILITY OF GRAPH CONVOLUTIONAL NEURAL NETWORKS THROUGH THE LENS OF SMALL PERTURBATION ANALYSIS
4369	STABLE DISTILLATION: REGULARIZING CONTINUED PRE-TRAINING FOR LOW-RESOURCE AUTOMATIC SPEECH RECOGNITION
1135	STABLE KNOWLEDGE TRANSFER FOR CONTRASTIVE DISTILLATION
1714	Stable Optimization for Large Vision Model Based Deep Image Prior in Cone-Beam CT Reconstruction
8327	STABLEMISS+: PREDICTION WITH INCOMPLETE DATA UNDER AGNOSTIC MASK DISTRIBUTION SHIFT
4813	STACK-AND-DELAY: A NEW CODEBOOK PATTERN FOR MUSIC GENERATION
8493	Stage-Regularized Neural Stein Critics for Testing Goodness-of-Fit of Generative Models
4085	STAR: DISTILLING SPEECH TEMPORAL RELATION FOR LIGHTWEIGHT SPEECH SELF-SUPERVISED LEARNING MODELS
8518	STATE-AUGMENTED INFORMATION ROUTING IN COMMUNICATION SYSTEMS WITH GRAPH NEURAL NETWORKS
7895	STATEFUL CONFORMER WITH CACHE-BASED INFERENCE FOR STREAMING AUTOMATIC SPEECH RECOGNITION
6877	Statistical and Computational Limits of Detecting and Recovering Hidden Submatrices
8646	STEALTHY BACKDOOR ATTACK TOWARDS FEDERATED AUTOMATIC SPEAKER VERIFICATION
8321	STEIN VARIATIONAL GRADIENT DESCENT-BASED DETECTION FOR RANDOM ACCESS WITH PREAMBLES IN MTC
7555	StemGen: A music generation model that listens
7128	STEREO-MATCHING KNOWLEDGE DISTILLED MONOCULAR DEPTH ESTIMATION FILTERED BY MULTIPLE DISPARITY CONSISTENCY

<b>Paper Number</b>	<b>Paper Title</b>
4783	STEREOPHONIC MUSIC SOURCE SEPARATION WITH SPATIALLY-INFORMED BRIDGING BAND-SPLIT NETWORK
9384	Stethoscope-guided Supervised Contrastive Learning for Cross-domain Adaptation on Respiratory Sound Classification
6533	Stochastic Configuration Networks for Laboratory Seismic Time-to-Failure Prediction
1116	STOFNET: SUPER-RESOLUTION TIME OF FLIGHT NETWORK
6910	STORYTTS: A HIGHLY EXPRESSIVE TEXT-TO-SPEECH DATASET WITH RICH TEXTUAL EXPRESSIVENESS ANNOTATIONS
9517	Straighforward adaptation of particle filter to fish eye images for top view pedestrian tracking
9061	STRATEGIC ARMS WITH SIDE COMMUNICATION PREVAIL OVER LOW-REGRET MAB ALGORITHMS
1032	STREAMING ACTIVE LEARNING FOR REGRESSION PROBLEMS USING REGRESSION VIA CLASSIFICATION
4397	STREAMING ANCHOR LOSS: AUGMENTING SUPERVISION WITH TEMPORAL SIGNIFICANCE
4781	StreamVC: Real-Time Low-Latency Voice Conversion
10217	STRING SOUND SYNTHESIZER ON GPU-ACCELERATED FINITE DIFFERENCE SCHEME
11528	STRONG LABELING OF SOUND EVENTS USING CROWDSOURCED WEAK LABELS AND ANNOTATOR COMPETENCE ESTIMATION
6875	Structure matters: analyzing videos via graph neural networks for social media platform attribution
1224	Structure-Aware In-Air Handwritten Text Recognition With Graph-Guided Cross-Modality Translator
6689	STRUCTURE-INFORMED POSITIONAL ENCODING FOR MUSIC GENERATION
6950	STS-CCL: Spatial-Temporal Synchronous Contextual Contrastive Learning for Urban Traffic Forecasting
7278	STUDY OF ABUSE DETECTION IN CONTINUOUS SPEECH FOR INDIAN LANGUAGES

<b>Paper Number</b>	<b>Paper Title</b>
1882	Style Adaptation for Domain-Adaptive Semantic Segmentation
8710	Style Factorization: Explore Diverse Style Variation for Domain Generalization
5611	STYLECAP: AUTOMATIC SPEAKING-STYLE CAPTIONING FROM SPEECH BASED ON SPEECH AND LANGUAGE SELF-SUPERVISED LEARNING MODELS
8611	STYLESPEECH: SELF-SUPERVISED STYLE ENHANCING WITH VQ-VAE-BASED PRE-TRAINING FOR EXPRESSIVE AUDIOBOOK SPEECH SYNTHESIS
11870	SUB-BAND AND FULL-BAND INTERACTIVE U-NET WITH DPRNN FOR DEMIXING CROSS-TALK STEREO MUSIC
2710	SUBDIVISION FEATURES-GUIDED BRAIN MRI SUPER-RESOLUTION VIA FORWARD AND BACKWARD PROPAGATION
8168	SUBGROUP IDENTIFICATION THROUGH MULTIPLEX COMMUNITY STRUCTURE WITHIN FUNCTIONAL CONNECTIVITY NETWORKS
7220	Subnetwork-to-go: Elastic Neural Network with Dynamic Training and Customizable Inference
3250	Subspace-Based Co-Array Processing for Nested Arrays Without Eigendecomposition
7881	SUBSPACE-BASED DETECTION IN OFDM ISAC SYSTEMS UNDER DIFFERENT CONSTELLATIONS
8695	SUBTYPE-SPECIFIC BIOMARKERS OF ALZHEIMER'S DISEASE FROM ANATOMICAL AND FUNCTIONAL CONNECTOMES VIA GRAPH NEURAL NETWORKS
6240	SUMMARIZING COMMUNITY-BASED QUESTION-ANSWER PAIRS WITH FOCUS RECTIFICATION
11947	SUMMARY ON THE MULTIMODAL INFORMATION-BASED SPEECH PROCESSING (MISP) 2023 CHALLENGE
6809	SUNFLOWER STRATEGY FOR BAYESIAN RELATIONAL DATA ANALYSIS
3594	SuperCodec: A Neural Speech Codec with Selective Back-Projection Network

<b>Paper Number</b>	<b>Paper Title</b>
11468	SUPERIORIZED ADAPTIVE PROJECTED SUBGRADIENT METHOD WITH APPLICATION TO MIMO DETECTION
4342	Supplementing Missing Visions via Dialog for Scene Graph Generations
5541	SURFACE-CONSTRAINED PROGRESSIVE FEATURE PRESERVING POINT CLOUD COMPRESSION
8943	SVAD: A ROBUST, LOW-POWER, AND LIGHT-WEIGHT VOICE ACTIVITY DETECTION WITH SPIKING NEURAL NETWORKS
1457	SWEEPMM: A HIGH-QUALITY MULTIMODAL DATASET FOR SWEEPING ROBOTS IN HOME SCENARIOS FOR VISION-LANGUAGE MODEL
6386	SYLLABLE LEVEL FEATURES FOR PARKINSON'S DISEASE DETECTION FROM SPEECH
1272	Symmetric Consistency with Cross-Domain Mixup for Cross-modality Cardiac Segmentation
8248	SYMMETRIC VAR(1) MODELLING WITH GUARANTEED STABILITY
6456	SYNCFUSION: MULTIMODAL ONSET-SYNCHRONIZED VIDEO-TO-AUDIO FOLEY SYNTHESIS
2154	SYNCHFORMER: EFFICIENT SYNCHRONIZATION FROM SPARSE CUES
9955	SYNONYM REPLACEMENT AND GENERATION ENHANCEMENT FOR DOCUMENT AUGMENTATION
2541	SYNTHE-SEES: FACE BASED TEXT-TO-SPEECH FOR VIRTUAL SPEAKER
6567	SYNTHESIZING A $\beta$ -PET VIA AN IMAGE AND LABEL CONDITIONING LATENT DIFFUSION MODEL FOR DETECTING AMYLOID STATUS
3018	SYNTHESIZING BLACK-BOX ANTI-FORENSICS DEEPPAKES WITH HIGH VISUAL QUALITY
3093	SYNTHETIC CONVERSATIONS IMPROVE MULTI-TALKER ASR
9082	Synthia's Melody: A Benchmark Framework for Unsupervised Domain Adaptation in Audio
8215	SYNTHTAB: LEVERAGING SYNTHESIZED DATA FOR GUITAR TABLATURE TRANSCRIPTION

<b>Paper Number</b>	<b>Paper Title</b>
6416	SYNVOX2: TOWARDS A PRIVACY-FRIENDLY VOXCELEB2 DATASET
9501	TA2P: Task-Aware Adaptive Pruning Method for Image Classification on Edge Devices
3915	TACKLING ELECTRODE SHIFT IN GESTURE RECOGNITION WITH HD-EMG ELECTRODE SUBSETS
1021	TACOS: LEARNING TEMPORALLY STRUCTURED EMBEDDINGS FOR FEW-SHOT KEYWORD SPOTTING WITH DYNAMIC TIME WARPING
10022	Tag Antenna Structure Calibrated Backscattering Signal Detection
3243	TAIL CLASSES MATTER: LONG-TAILED OBJECT DETECTION REVISITED
6614	TALDS-Net: Task-Aware Adaptive Local Descriptors Selection for Few-shot Image Classification
9657	TALKING FACE GENERATION FOR IMPRESSION CONVERSION CONSIDERING SPEECH SEMANTICS
9571	TALKNCE: IMPROVING ACTIVE SPEAKER DETECTION WITH TALK-AWARE CONTRASTIVE LEARNING
1309	TAMING PROMPT-BASED DATA AUGMENTATION FOR LONG-TAILED EXTREME MULTI-LABEL TEXT CLASSIFICATION
7417	TARGET LOCALIZATION BASED ON MULTISTATIC MIMO RADAR VIA DOUBLE COUPLED CANONICAL POLYADIC DECOMPOSITION
5824	TARGET OPTIMIZATION DIRECTION GUIDED TRANSFER LEARNING FOR IMAGE
10171	TARGET SIGNAL POWER IMPROVEMENT AND CLUTTER SUPPRESSION VIA BEAMFORMING FOR INTEGRATED SENSING AND COMMUNICATION SYSTEMS
3147	TARGET SPEAKER EXTRACTION BY DIRECTLY EXPLOITING CONTEXTUAL INFORMATION IN THE TIME-FREQUENCY DOMAIN
2997	TARGET SPEECH EXTRACTION WITH PRE-TRAINED SELF-SUPERVISED LEARNING MODELS

<b>Paper Number</b>	<b>Paper Title</b>
7921	TAROT: A Hierarchical Framework with Multitask Co-Pretraining on Semi-Structured Data towards Effective Person-Job Fit
5684	TASK INDICATING TRANSFORMER FOR TASK-CONDITIONAL DENSE PREDICTIONS
7382	TASK ORIENTED DIALOGUE AS A CATALYST FOR SELF-SUPERVISED AUTOMATIC SPEECH RECOGNITION
7990	TASK SELECTION AND ASSIGNMENT FOR MULTI-MODAL MULTI-TASK DIALOGUE ACT CLASSIFICATION WITH NON-STATIONARY MULTI-ARMED BANDITS
8428	Task vector algebra for ASR models
5101	TASK-WISE PROMPT QUERY FUNCTION FOR REHEARSAL-FREE CONTINUAL LEARNING
2452	TB-RESNET: BRIDGING THE GAP FROM TDNN TO RESNET IN AUTOMATIC SPEAKER VERIFICATION WITH TEMPORAL-BOTTLENECK ENHANCEMENT
2891	TCMP: END-TO-END TOPOLOGICALLY CONSISTENT MAGNITUDE PRUNING FOR MINIATURIZED GRAPH CONVOLUTIONAL NETWORKS
3265	TCNAS: TRANSFORMER ARCHITECTURE EVOLVING IN CODE CLONE DETECTION
10270	TD-GPT:TARGET PROTEIN-SPECIFIC DRUG MOLECULE GENERATION GPT
6073	TDT-KWS: FAST AND ACCURATE KEYWORD SPOTTING USING TOKEN-AND-DURATION TRANSDUCER
4775	TEMPLATE-GUIDED DATA AUGMENTATION FOR UNBIASED SCENE GRAPH GENERATION
8950	TEMPO ESTIMATION AS FULLY SELF-SUPERVISED BINARY CLASSIFICATION
2313	TEMPORAL CONDITIONAL CODING FOR DYNAMIC POINT CLOUD GEOMETRY COMPRESSION
2042	TEMPORAL CONVOLUTION SHRINKAGE NETWORK FOR KEYWORD SPOTTING
6927	TEMPORAL INCONSISTENCY-BASED ACTIVE LEARNING
1871	TEMPORAL KNOWLEDGE GRAPH EMBEDDING USING HOUSEHOLDER TRANSFORMATIONS



<b>Paper Number</b>	<b>Paper Title</b>
5718	TEMPORAL RELATIONAL CONTEXT LEARNING FOR EXTRAPOLATION REASONING ON TEMPORAL KNOWLEDGE GRAPHS
7568	TEMPORALLY-GUIDED TOTAL VARIATION FOR ROBUST SPATIOTEMPORAL FUSION OF SATELLITE IMAGES
4063	Temporal-Spatial Prediction: pre-training on diverse datasets for EEG classification
7341	T-ENFP: AN EFFICIENT TRANSFORMER ENCODER-BASED SYSTEM FOR DRIVING BEHAVIOR CLASSIFICATION
8104	TEN-GUARD: TENSOR DECOMPOSITION FOR BACKDOOR ATTACK DETECTION IN DEEP NEURAL NETWORKS
5843	Tensor decomposition-based data fusion for biomarker extraction from multiple EEG experiments
9792	Tensor Graph Decomposition for Temporal Networks
3982	Tensor Low-rank Approximation of Finite-horizon Value Functions
2310	TENSOR RECONSTRUCTION-BASED SPARSE ARRAY 2-D DOA ESTIMATION OF MIXED COHERENT AND UNCORRELATED SIGNALS
8579	TENSOR-GUIDED INTERPOLATION FOR OFF-GRID POWER SPECTRUM MAP CONSTRUCTION
2800	TENSORIAL CONVOLUTIVE BLIND SOURCE SEPARATION
9775	Test-Time Distribution Learning Adapter For Cross-Modal Visual Reasoning
10293	Text Region Multiple Information Perception Network for Scene Text Detection
1797	TEXT2AVATAR: TEXT TO 3D HUMAN AVATAR GENERATION WITH CODEBOOK-DRIVEN BODY CONTROLLABLE ATTRIBUTE
2059	Text-Driven 3D Human Generation via 2D Image Collections
4720	TEXT-DRIVEN TALKING FACE SYNTHESIS BY REPROGRAMMING AUDIO-DRIVEN MODELS
5094	TEXT-ONLY UNSUPERVISED DOMAIN ADAPTATION FOR NEURAL TRANSDUCER-BASED ASR PERSONALIZATION USING SYNTHESIZED DATA

<b>Paper Number</b>	<b>Paper Title</b>
2483	TEXTROLSPEECH: A TEXT STYLE CONTROL SPEECH CORPUS WITH CODEC LANGUAGE TEXT-TO-SPEECH MODELS
4034	TEXTUAL TOKENS CLASSIFICATION FOR MULTI-MODAL ALIGNMENT IN VISION-LANGUAGE TRACKING
4348	Texture and normal map estimation for 3D face reconstruction
7196	Texture-Unet: A Texture-Aware Network for Bone Marrow Smear Whole-slide Image Region of Interest Segmentation
2570	TEXT-VIDEO COMPLETION NETWORKS WITH MOTION COMPENSATION AND ATTENTION AGGREGATION
7320	T-FOLEY: A CONTROLLABLE WAVEFORM-DOMAIN DIFFUSION MODEL FOR TEMPORAL-EVENT-GUIDED FOLEY SOUND SYNTHESIS
5081	TF-SepNet: An Efficient 1D Kernel Design in CNNs for Low-Complexity Acoustic Scene Classification
6958	THE 2ND CLARITY PREDICTION CHALLENGE: A MACHINE LEARNING CHALLENGE FOR HEARING AID INTELLIGIBILITY PREDICTION
11948	THE 2ND E-PREVENTION CHALLENGE: PSYCHOTIC AND NON-PSYCHOTIC RELAPSE DETECTION USING WEARABLE-BASED DIGITAL PHENOTYPING
9320	THE COLLABORATION OF 3D CONVOLUTIONS AND CRO-TSM IN LIPREADING
11889	The Data-Driven Radio Frequency Signal Separation Challenge
8694	The Devil is in Details: Delving into Lite FFN Design for Vision Transformers
8459	THE DOUBLE-EDGED SWORD OF AI SAFETY: BALANCING ANOMALY DETECTION AND OOD GENERALIZATION VIA MODEL ANCHORING
7679	THE EFFECTS OF LOUDNESS AND SMILING ON TIMBRE FEATURES: IMPLICATIONS FOR CHARISMATIC VOICES IN MANDARIN, GERMAN AND DANISH
11869	THE FAWAISPEECH SYSTEM FOR MULTI-CHANNEL SPEECH RECOGNITION IN ICMC-ASR CHALLENGE

<b>Paper Number</b>	<b>Paper Title</b>
11856	THE FOSA FER SYSTEM FOR THE ICASSP2024 IN-CAR MULTI-CHANNEL AUTOMATIC SPEECH RECOGNITION CHALLENGE
11881	THE ICASSP 2024 AUDIO DEEP PACKET LOSS CONCEALMENT GRAND CHALLENGE
11918	The ICASSP SP Cadenza Challenge: Music Demixing/Remixing For Hearing Aids
2633	THE JOINT GRID-FREE DOA AND POLARIZATION ESTIMATION ALGORITHM BASED ON ATOMIC NORM MINIMIZATION
9142	THE MULTIMODAL INFORMATION BASED SPEECH PROCESSING (MISP) 2023 CHALLENGE: AUDIO-VISUAL TARGET SPEAKER EXTRACTION
8139	The power of few: accelerating and enhancing data reweighting with coreset selection
8022	THE RAO, WALD, AND LIKELIHOOD-RATIO TESTS UNDER GENERALIZED SELF-CONCORDANCE
11851	THE ROYALFLUSH AUTOMATIC SPEECH DIARIZATION AND RECOGNITION SYSTEM FOR IN-CAR MULTI-CHANNEL AUTOMATIC SPEECH RECOGNITION CHALLENGE
3181	THE SELECTIVITY AND COMPETITION OF THE MIND'S EYE IN VISUAL PERCEPTION
11905	THE THU-HCSI MULTI-SPEAKER MULTI-LINGUAL FEW-SHOT VOICE CLONING SYSTEM FOR LIMMITS'24 CHALLENGE
11893	THE USTC SYSTEM FOR CADENZA 2024 CHALLENGE
11852	THE USTC-NERCSLIP SYSTEMS FOR THE ICMC-ASR CHALLENGE
11880	The XMUSPEECH SYSTEM FOR AUDIO-VISUAL TARGET SPEAKER EXTRACTION IN MISP 2023 CHALLENGE
8924	Theme-enhanced Hard Negative Sample Mining for Open-domain Question Answering
10373	Think as People: Context-driven Multi-image News Captioning with Adaptive Dual Attention
11519	Third-Order Nested Array: An Optimal Geometry For Third-Order Cumulants Based Array Processing

<b>Paper Number</b>	<b>Paper Title</b>
7463	THREE-DIMENSIONAL DECOUPLED ATOMIC NORM MINIMIZATION
7251	THREE-DIMENSIONAL SOUND WAVE PROPAGATION REPRODUCTION BY CE-FDTD SIMULATION APPLYING ACTUAL RADIATION CHARACTERISTICS
2839	Three-dimensional Spatial-Temporal Near-Field Passive Localization Based on an Exact Spatial Propagation Model
4008	Through-the-Wall Radar Imaging with wall clutter removal via Riemannian optimization on the fixed-rank manifold
7428	TIA: A TEACHING INTONATION ASSESSMENT DATASET IN REAL TEACHING SITUATIONS
8273	TIMBRE-TRAP: A LOW-RESOURCE FRAMEWORK FOR INSTRUMENT-AGNOSTIC MUSIC TRANSCRIPTION
5511	Time Changed Normalizing Flows for accurate SDE modeling
2943	TIME-INTERVAL VISUAL SALIENCY PREDICTION IN MAMMOGRAM READING
4310	TIME-MODULATED INTELLIGENT REFLECTING SURFACE FOR WAVEFORM SECURITY
4611	TITAN: BRINGING THE DEEP IMAGE PRIOR TO IMPLICIT REPRESENTATIONS
7149	TNFORMER: SINGLE-PASS MULTILINGUAL TEXT NORMALIZATION WITH A TRANSFORMER DECODER MODEL
2159	TODM: TRAIN ONCE DEPLOY MANY EFFICIENT SUPERNET-BASED RNN-T COMPRESSION FOR ON-DEVICE ASR MODELS
1875	TOKEN-BASED SPATIOTEMPORAL REPRESENTATION OF THE EVENTS
2160	TokenMotion: Motion-Guided Vision Transformer for Video Camouflaged Object Detection Via Learnable Token Selection
3739	TOPOLOGICAL NEURAL NETWORKS OVER THE AIR
3872	TOPOLOGY-DEPENDENT PRIVACY BOUND FOR DECENTRALIZED FEDERATED LEARNING

<b>Paper Number</b>	<b>Paper Title</b>
8458	Topology-Regularized Self-Knowledge Distillation for Transductive-Inductive Learning of Brain Disorder Diagnosis
1979	Touring sampling with pushforward maps
1145	Toward Quantifiable Face Age Transformation
3461	TOWARD SUFFICIENT SPATIAL-FREQUENCY INTERACTION FOR GRADIENT-AWARE UNDERWATER IMAGE ENHANCEMENT
9139	TOWARDS 3D COMPUTATIONAL PERSICOPY WITH AN ORDINARY CAMERA: A SEPARABLE NON-LINEAR LEAST SQUARES FORMULATION.
2292	TOWARDS A UNIFIED VIEW OF ADVERSARIAL TRAINING: A CONTRASTIVE PERSPECTIVE
4361	Towards a World-English Language Model for On-Device Virtual Assistants
8833	TOWARDS AN INTERPRETABLE REPRESENTATION OF SPEAKER IDENTITY VIA PERCEPTUAL VOICE QUALITIES
7249	TOWARDS AN OBJECTIVE QUALITY METRIC FOR INTERPOLATED DIRECTIONAL ROOM IMPULSE RESPONSES
1856	Towards ASR Robust Spoken Language Understanding Through In-Context Learning With Word Confusion Networks
3589	TOWARDS AUTOMATIC DATA AUGMENTATION FOR DISORDERED SPEECH RECOGNITION
7617	Towards Building the FederatedGPT: Federated Instruction Tuning
1027	TOWARDS CONTROLLED TABLE-TO-TEXT GENERATION WITH SCIENTIFIC REASONING
9383	TOWARDS DISEASE-AWARE SELF-SUPERVISED DYNAMIC BRAIN NETWORK LEARNING FOR MENTAL DIAGNOSIS
2336	Towards Efficient Modeling and Inference in Multi-Dimensional Gaussian Process State-Space Models
9901	TOWARDS ENABLING DPOAE ESTIMATION ON SINGLE-SPEAKER EARBUDS

<b>Paper Number</b>	<b>Paper Title</b>
4314	TOWARDS END-TO-END SPOKEN GRAMMATICAL ERROR CORRECTION
7014	TOWARDS FASTER END-TO-END DATA TRANSMISSION OVER VOICE CHANNELS
2360	Towards Generic Deepfake Detection with Dynamic Curriculum
8511	TOWARDS HIGH RESOLUTION WEATHER MONITORING WITH SOUND DATA
4710	TOWARDS HIGH-PERFORMANCE AND LOW-LATENCY FEATURE-BASED SPEAKER ADAPTATION OF CONFORMER SPEECH RECOGNITION SYSTEMS
3631	TOWARDS IMPROVING SPEECH EMOTION RECOGNITION USING SYNTHETIC DATA AUGMENTATION FROM EMOTION CONVERSION
6437	TOWARDS INTELLIGENT DESIGN: A SELF-DRIVEN FRAMEWORK FOR COLLOCATED CLOTHING SYNTHESIS LEVERAGING FASHION STYLES AND TEXTURES
9502	TOWARDS INTERPRETABILITY OF AUTOMATIC PHONEME ANALYSIS IN CLEFT LIP AND PALATE SPEECH
7040	Towards Multi-domain Face Landmark Detection with Synthetic Data from Diffusion Model
8875	TOWARDS OMNISCIENT FEATURE ALIGNMENT FOR VIDEO RESCALING
4674	TOWARDS OPTIMAL VOICE DISENTANGLEMENT WITH WEAK SUPERVISION
4120	TOWARDS OPTIMIZED MULTI-CHANNEL MODULO-ADCS: MODULI SELECTION STRATEGIES AND BIT DEPTH ANALYSIS
3107	TOWARDS PRACTICAL AND EFFICIENT IMAGE-TO-SPEECH CAPTIONING WITH VISION-LANGUAGE PRE-TRAINING AND MULTI-MODAL TOKENS
2683	TOWARDS RESOURCE-EFFICIENT AND SECURE FEDERATED MULTIMEDIA RECOMMENDATION
4755	TOWARDS ROBUST MULTIMODAL PROMPTING WITH MISSING MODALITIES

<b>Paper Number</b>	<b>Paper Title</b>
2885	TOWARDS UNIVERSAL SPEECH DISCRETE TOKENS: A CASE STUDY FOR ASR AND TTS
6022	TOWARDS VIDEO-TEXT RETRIEVAL ADVERSARIAL ATTACK
3370	T-Pixel2Mesh: Combining Global and Local Transformer for 3D Mesh Generation from a Single Image
1725	TRACKING BEYOND THE UNAMBIGUOUS RANGE WITH MODULO SINGLE-PHOTON LIDAR
7042	TRACKING OF MULTIPLE SPAWNING TARGETS WITH HETEROGENEOUS SENSORS FOR SEABED-TO-SPACE SITUATIONAL AWARENESS
4675	TraDeS++: Enhancing Multi-Object Tracking of Real Low Confidence Targets Using a Pyramid-like Self-Attention Model
4360	Train Long and Test Long: Leveraging Full Document Contexts in Speech Processing
10441	Training a Radial Basis Function Network Under Transformed Probability Measure
2155	TRAINING AUDIO CAPTIONING MODELS WITHOUT AUDIO
9291	TRAINING GENERATIVE ADVERSARIAL NETWORK-BASED VOCODER WITH LIMITED DATA USING AUGMENTATION-CONDITIONAL DISCRIMINATOR
4206	Training Ultra-Low-Latency Spiking Neural Networks from Scratch
4327	TRAJECTORY SET EMPOWERED HYPERGRAPH TRANSFORMER FOR MOBILE SENSOR BASED TRAFFIC PREDICTION
1271	TRANSAVS: END-TO-END AUDIO-VISUAL SEGMENTATION WITH TRANSFORMER
3828	TransCycle: A Data Augmentation Method For 3D Human Pose Estimation
7862	Transducers with Pronunciation-aware Embeddings for Automatic Speech Recognition
9891	TRANSENTENCE: SPEECH-TO-SPEECH TRANSLATION VIA LANGUAGE-AGNOSTIC SENTENCE-LEVEL SPEECH ENCODING WITHOUT LANGUAGE-PARALLEL DATA

<b>Paper Number</b>	<b>Paper Title</b>
9048	TRANSFER THE LINGUISTIC REPRESENTATIONS FROM TTS TO ACCENT CONVERSION WITH NON-PARALLEL DATA
8649	Transferable Models for Bioacoustics with Human Language Supervision
4183	Transferring Structure Knowledge: A New Task to Fake news Detection Towards Cold-Start Propagation
2468	TRANSFORMER MODEL WITH MULTI-TYPE CLASSIFICATION DECISIONS FOR INTRUSION ATTACK DETECTION OF TRACK TRAFFIC AND VEHICLE
6825	TRANSFORMER-INSPIRED LIGHTWEIGHT MODEL FOR EFFICIENT TIME SERIES FORECASTING
7384	TRANSFORMING CARDIOVASCULAR HEALTH: A TRANSFORMER-BASED APPROACH TO CONTINUOUS, NON-INVASIVE BLOOD PRESSURE ESTIMATION VIA RADAR SENSING
3880	TRANSLATOTRON 3: SPEECH TO SPEECH TRANSLATION WITH MONOLINGUAL DATA
6021	TRANSMIT BEAMPATTERN OPTIMIZATION FOR MIMO-ISAC SYSTEMS WITH HYBRID BEAMFORMING
9741	TRANSMITTING DATA THROUGH RECONFIGURABLE INTELLIGENT SURFACE: A SPATIAL SIGMA-DELTA MODULATION APPROACH
3480	TRANSMUSIC: A TRANSFORMER-AIDED SUBSPACE METHOD FOR DOA ESTIMATION WITH LOW-RESOLUTION ADCS
4296	TREE NETWORK DESIGN FOR FASTER DISTRIBUTED MACHINE LEARNING PROCESS WITH DISTRIBUTED DUAL COORDINATE ASCENT
9965	TREE OF UNCERTAIN THOUGHTS REASONING FOR LARGE LANGUAGE MODELS
9280	TREEMIL: A MULTI-INSTANCE LEARNING FRAMEWORK FOR TIME SERIES ANOMALY DETECTION WITH INEXACT SUPERVISION
1496	TREND-HEURISTIC REINFORCEMENT LEARNING FRAMEWORK FOR NEWS-ORIENTED STOCK PORTFOLIO MANAGEMENT



<b>Paper Number</b>	<b>Paper Title</b>
3406	TRET: Two Stream-based Regionally Enhanced Transformers for Person Re-identification
1761	TRLS: A Time Series Representation Learning Framework via Spectrogram for Medical Signal Processing
2609	TRUSTED DEEP DOMAIN ADAPTATION WITH UNCERTAINTY MEASURE BASED ON EVIDENCE THEORY
5312	TRUST-SER: ON THE TRUSTWORTHINESS OF FINE-TUNING PRE-TRAINED SPEECH EMBEDDINGS FOR SPEECH EMOTION RECOGNITION
6926	T-SOT FNT: STREAMING MULTI-TALKER ASR WITH TEXT-ONLY DOMAIN ADAPTATION CAPABILITY
8109	TURN-TAKING AND BACKCHANNEL PREDICTION WITH ACOUSTIC AND LARGE LANGUAGE MODEL FUSION
7797	TWO-EDGE-RESOLVED 3D NON-LINE-OF-SIGHT IMAGING: A FISHER INFORMATION EQUALIZED DISCRETIZATION
3833	TWO-STAGE ACOUSTIC ECHO CANCELLATION NETWORK WITH DUAL-PATH ALIGNMENT
11861	TWO-STAGE NEURAL NETWORK MODEL WITH PACKET LOSS DETECTION FOR ICASSP 2024 PLC CHALLENGE
6370	TWO-STAGE TRANSFER LEARNING FOR FUSION AND CLASSIFICATION OF AIRBORNE HYPERSPECTRAL IMAGERY
1997	TWO-STEP KNOWLEDGE DISTILLATION FOR TINY SPEECH ENHANCEMENT
4890	TYPE-AWARE DECODING VIA EXPLICITLY AGGREGATING EVENT INFORMATION FOR DOCUMENT-LEVEL EVENT EXTRACTION
4081	U2R: UNDERWATER ULTRASONIC REFLECTION WAVE DATASET TOWARD POSE-INVARIANT MATERIAL RECOGNITION
2477	UAMIX-MAE: EFFICIENT TUNING OF PRETRAINED AUDIO TRANSFORMERS WITH UNSUPERVISED AUDIO MIXTURES
4072	UAV Operation Time Minimization for Wireless-Powered Data Collection
8395	UAV-based Dynamic Object Tracking with Radio Map

<b>Paper Number</b>	<b>Paper Title</b>
2963	ULTRA LOW COMPLEXITY DEEP LEARNING BASED NOISE SUPPRESSION
1760	ULTRA-LIGHTWEIGHT NEURAL DIFFERENTIAL DSP VOCODER FOR HIGH QUALITY SPEECH SYNTHESIS
4785	ULTRA-LOW DELAY LOSSLESS COMPRESSION OF HIGHER ORDER AMBISONICS
2728	UNAD: UNIVERSAL ANATOMY-INITIALIZED NOISE DISTRIBUTION LEARNING FRAMEWORK TOWARDS LOW-DOSE CT DENOISING
5126	UNCERTAINTY QUANTIFICATION IN DEEP LEARNING BASED KALMAN FILTERS
7672	UNCERTAINTY-GUIDED CONTRASTIVE LEARNING FOR SINGLE SOURCE DOMAIN GENERALISATION
7274	Uncertainty-guided Person Search model with Auxiliary Shallow Feature Exploration
8637	Uncertainty-Guided Physics-Driven Deep Learning Reconstruction via Cyclic Measurement Consistency
2884	UNCOVERING STRONG TIES: A STUDY OF INDIRECT SYBIL ATTACK ON SIGNED SOCIAL NETWORK
3375	UNDERLYING-COMPLEMENTARITY AND SURROUNDING-CORRESPONDENCE FOR MULTI-VIEW CLUSTERING
7151	UNDERSTANDING DATA AUGMENTATION FROM A ROBUSTNESS PERSPECTIVE
6550	Understanding Gaussian Noise Mismatch: A Hellinger Distance Approach
3047	UNDERSTANDING PROBE BEHAVIORS THROUGH VARIATIONAL BOUNDS OF MUTUAL INFORMATION
8479	UNeC: UNSUPERVISED EXPLORING IN CONTROLLABLE SPACE
8234	UNIDEAL: CURRICULUM KNOWLEDGE DISTILLATION FEDERATED LEARNING
4599	UNIDIRECTIONAL BRAIN-COMPUTER INTERFACE: ARTIFICIAL NEURAL NETWORK ENCODING NATURAL IMAGES TO fMRI RESPONSE IN THE VISUAL CORTEX
4609	Unified Analysis of Correlation-Aware Joint Sparse Support Recovery with $l_0$ -Norm Constraint

<b>Paper Number</b>	<b>Paper Title</b>
1838	UNIFIED PRETRAINING TARGET BASED VIDEO-MUSIC RETRIEVAL WITH MUSIC RHYTHM AND VIDEO OPTICAL FLOW INFORMATION
3197	UNIFIED PROBABILITY DISTRIBUTIONS OF GENERALIZED COMPOSITE FADING WITH INVERSE-TYPE DISTRIBUTIONS OF LARGE-SCALE SHADOWING/FLUCTUATIONS
7442	UNIFIED SPEECH AND GESTURE SYNTHESIS USING FLOW MATCHING
1421	Unified sRGB Real Noise Synthesizing with Adaptive Feature Modulation
5098	Unifying One-Shot Voice Conversion and Cloning with Disentangled Speech Representations
3504	UNIMODAL AGGREGATION FOR CTC-BASED SPEECH RECOGNITION
4238	UNINTENDED MEMORIZATION IN LARGE ASR MODELS, AND HOW TO MITIGATE IT
5564	UNITARY APPROXIMATE MESSAGE PASSING FOR MATRIX FACTORIZATION
8598	UNIT-DSR: DYSARTHIC SPEECH RECONSTRUCTION SYSTEM USING SPEECH UNIT NORMALIZATION
8854	UNIVERSAL ADVERSARIAL ATTACK AGAINST SPEAKER RECOGNITION MODELS
7795	UNIX-Encoder: A Universal X-Channel Speech Encoder for Ad-Hoc Microphone Array Speech Processing
5068	Unlabelled Sensing with Priors: Algorithm and Bounds
2057	UNLEASHING TRIGGER-FREE EVENT DETECTION: REVEALING EVENT CORRELATIONS VIA A CONTRASTIVE DERANGEMENT FRAMEWORK
9159	UNLOCKING DEEP LEARNING: A BP-FREE APPROACH FOR PARALLEL BLOCK-WISE TRAINING OF NEURAL NETWORKS
2415	UNRAVEL ANOMALIES: AN END-TO-END SEASONAL-TREND DECOMPOSITION APPROACH FOR TIME SERIES ANOMALY DETECTION
5892	UNRAVELING EXPLAINABLE REINFORCEMENT LEARNING USING BEHAVIOR TREE STRUCTURES

<b>Paper Number</b>	<b>Paper Title</b>
7152	UNRESTRICTED GLOBAL-PHASE-BIAS AWARE SINGLE-CHANNEL SPEECH ENHANCEMENT WITH CONFORMER-BASED METRIC GAN
9684	UNROLLED PROXIMAL GRADIENT DESCENT METHOD FOR NON-NEGATIVE LEAST SQUARES PROBLEM
2242	UNSUPERVISED ACCENT ADAPTATION THROUGH MASKED LANGUAGE MODEL CORRECTION OF DISCRETE SELF-SUPERVISED SPEECH UNITS
2311	UNSUPERVISED ACOUSTIC SCENE MAPPING BASED ON ACOUSTIC FEATURES AND DIMENSIONALITY REDUCTION
6532	UNSUPERVISED ANOMALY DETECTION FOR MULTIVARIATE TIME SERIES USING DIFFUSION MODEL
1115	UNSUPERVISED CONTINUAL LEARNING OF IMAGE REPRESENTATION VIA REMEMORY-BASED SIMSIAM
1219	UNSUPERVISED DISPARITY ESTIMATION FOR LIGHT FIELD VIDEOS
8874	UNSUPERVISED EXTRACTIVE DIALOGUE SUMMARIZATION IN HYPERDIMENSIONAL SPACE
7726	UNSUPERVISED HARMONIC PARAMETER ESTIMATION USING DIFFERENTIABLE DSP AND SPECTRAL OPTIMAL TRANSPORT
5007	UNSUPERVISED HUMAN ACTIVITY RECOGNITION VIA LARGE LANGUAGE MODELS AND ITERATIVE EVOLUTION
2887	UNSUPERVISED LEARNING BASED END-TO-END DELAYLESS GENERATIVE FIXED-FILTER ACTIVE NOISE CONTROL
5764	UNSUPERVISED LEARNING OF FACIAL OPTICAL FLOW VIA OCCLUSION-AWARE GLOBAL-LOCAL MATCHING
1292	UNSUPERVISED LEARNING OF NEURAL SEMANTIC MAPPINGS WITH THE HUNGARIAN ALGORITHM FOR COMPOSITIONAL SEMANTICS
4323	UNSUPERVISED MULTI-CHANNEL SEPARATION AND ADAPTATION
4040	UNSUPERVISED MULTI-DOMAIN DATA SELECTION FOR ASR FINE-TUNING

<b>Paper Number</b>	<b>Paper Title</b>
7324	Unsupervised Multiple Choices Question Answering via Universal Corpus
7920	UNSUPERVISED MULTIPLE DOMAIN TRANSLATION THROUGH CONTROLLED DISENTANGLEMENT IN VARIATIONAL AUTOENCODER
7497	UNSUPERVISED OPTIMAL POWER FLOW USING GRAPH NEURAL NETWORKS
7192	UNSUPERVISED PITCH-TIMBRE DISENTANGLEMENT OF MUSICAL INSTRUMENTS USING A JACOBIAN DISENTANGLED SEQUENTIAL AUTOENCODER
11885	UNSUPERVISED RELAPSE DETECTION USING WEARABLE-BASED DIGITAL PHENOTYPING FOR THE 2ND E-PREVENTION CHALLENGE
8620	UNSUPERVISED REMOTE SENSING HAZE REMOVAL BASED ON SALIENCY-GUIDED TRANSMISSION REFINEMENT
9017	Unsupervised Speech Enhancement with Diffusion-based Generative Models
4604	UNSUPERVISED SPEECH RECOGNITION WITH N-SKIPGRAM AND POSITIONAL UNIGRAM MATCHING
5793	UNSUPERVISED TOPIC-CONDITIONAL EXTRACTIVE SUMMARIZATION
7532	UPDATED CORPORA AND BENCHMARKS FOR LONG-FORM SPEECH RECOGNITION
9637	UPLINK SYMBOL DETECTION IN DYNAMIC TDD MIMO SYSTEMS WITH AP-AP INTERFERENCE
2781	URBAN TRAFFIC FLOW FORECASTING BASED ON SPATIAL-TEMPORAL GRAPH CONTRASTIVE LEARNING
8183	USEE: UNIFIED SPEECH ENHANCEMENT AND EDITING WITH CONDITIONAL DIFFUSION MODELS
3319	USER-ASSISTED NETWORKED SENSING IN OFDM CELLULAR NETWORK WITH ERRONEOUS ANCHOR POSITION INFORMATION
7047	USING CLUSTERING TO IMPROVE THE PERFORMANCE OF FEW-SHOT LEARNING
1628	Using Temporal Consistency for Compressed Sensing in High-Resolution mmWave Sounding

<b>Paper Number</b>	<b>Paper Title</b>
4257	USM-Lite: Quantization and Sparsity Aware Fine-tuning for Speech Recognition with Universal Speech Models
7380	USM-SCD: MULTILINGUAL SPEAKER CHANGE DETECTION BASED ON LARGE PRETRAINED FOUNDATION MODELS
11490	Utility-driven Joint Caching and Bitrate Allocation for Real-Time Immersive Videos
8342	UTILIZING SECOND-ORDER INFORMATION IN NOISY INFORMATION-SHARING ENVIRONMENTS FOR DISTRIBUTED OPTIMIZATION
11452	Variable-Wise Diagonal Preconditioning for Primal-Dual Splitting: Design and Applications
7846	Variance Reduction Can Improve Trade-off In Multi-Objective Learning
7744	VARIATIONAL ANALYSIS OF ADVERSARIAL REGULARIZATION FOR SOLVING INVERSE PROBLEMS
6016	VARIATIONAL CONNECTIONIST TEMPORAL CLASSIFICATION FOR ORDER-PRESERVING SEQUENCE MODELING
7864	VCD: A Video Conferencing Dataset for Video Compression
9250	V-DDPM: MRI Rician Noise Removal Model Based on VST and DDPM
7681	VECTOR APPROXIMATE MESSAGE PASSING FOR NOT SO LARGE N.I.I.D. GENERALIZED I/O LINEAR MODELS
6422	VECTOR APPROXIMATE MESSAGE PASSING WITH ARBITRARY I.I.D. NOISE PRIORS
2667	VECTOR NONLINEAR HAWKES MODEL WITH INHIBITION
9008	VECTOR QUANTIZATION KNOWLEDGE TRANSFER FOR END-TO-END TEXT IMAGE MACHINE TRANSLATION
8172	VFD-NET: VOCODER FINGERPRINTS DETECTION FOR FAKE AUDIO
1688	VGDIFZERO: TEXT-TO-IMAGE DIFFUSION MODELS CAN BE ZERO-SHOT VISUAL GROUNDERS
8284	VIC-KD: VARIANCE-INVARIANCE-COVARIANCE KNOWLEDGE DISTILLATION TO MAKE KEYWORD SPOTTING MORE ROBUST AGAINST ADVERSARIAL ATTACKS

<b>Paper Number</b>	<b>Paper Title</b>
7215	VIDEO ANOMALY PREDICTION: PROBLEM, DATASET AND METHOD
3720	Video-language Graph Convolutional Network for Human Action Recognition
2274	View Crafting for Instance-Level Representation from Scene Images
3160	VIEWING WRITING AS VIDEO: OPTICAL FLOW BASED MULTI-MODAL HANDWRITTEN MATHEMATICAL EXPRESSION RECOGNITION
6122	VILAS: EXPLORING THE EFFECTS OF VISION AND LANGUAGE CONTEXT IN AUTOMATIC SPEECH RECOGNITION
11554	VIRTUAL BASS ENHANCEMENT VIA MUSIC DEMIXING
11888	Vision Transformer MST++: Efficient Hyperspectral Skin Reconstruction
9808	VISION TRANSFORMER WITH 2D EXPLICIT POSITION ENCODING
5048	VISION-SENSOR ATTENTION BASED CONTINUAL MULTIMODAL EGOCENTRIC ACTIVITY RECOGNITION
2752	VISUAL ADAPT FOR RGBD TRACKING
1891	VISUAL PROMPT TUNING FOR WEAKLY SUPERVISED PHRASE GROUNDING
3123	VISUAL SPEECH RECOGNITION FOR LANGUAGES WITH LIMITED LABELED DATA USING AUTOMATIC LABELS FROM WHISPER
5065	VISUAL-LINGUISTIC REPRESENTATION LEARNING WITH DEEP CROSS-MODALITY FUSION FOR REFERRING MULTI-OBJECT TRACKING
2640	VISUALLY DEHALLUCINATIVE INSTRUCTION GENERATION
3581	Visually Guided Binaural Audio Generation with Cross-modal Consistency
1530	VK-G2T: VISION AND CONTEXT KNOWLEDGE ENHANCED GLOSS2TEXT
7024	VL-FAS: DOMAIN GENERALIZATION VIA VISION-LANGUAGE MODEL FOR FACE ANTI-SPOOFING

<b>Paper Number</b>	<b>Paper Title</b>
3892	VMCC-NET: UNCOVERING CHALLENGING REGIONS IN SEMI-SUPERVISED MEDICAL IMAGE SEGMENTATION WITH VOXEL MASK BASED CYCLIC-CONSISTENCY NETWORK
1836	VOCAL FOLD DYNAMICS FOR AUTOMATIC DETECTION OF AMYOTROPHIC LATERAL SCLEROSIS FROM VOICE
7164	VOICE ANONYMIZATION FOR ALL - BIAS EVALUATION OF THE VOICE PRIVACY CHALLENGE BASELINE SYSTEMS
2004	VOICE TOXICITY DETECTION USING MULTI-TASK LEARNING
5036	VOICEFLOW: EFFICIENT TEXT-TO-SPEECH WITH RECTIFIED FLOW MATCHING
9312	VoiceLDM: Text-to-Speech with Environmental Context
7606	VOLUMETRIC 3D POINT CLOUD ATTRIBUTE COMPRESSION: LEARNED POLYNOMIAL BILATERAL FILTER FOR PREDICTION
2345	VoxBlink: A Large Scale Speaker Verification Dataset on Camera
9267	VOXMM: RICH TRANSCRIPTION OF CONVERSATIONS IN THE WILD
7874	VOXTLM: UNIFIED DECODER-ONLY MODELS FOR CONSOLIDATING SPEECH RECOGNITION, SYNTHESIS AND SPEECH, TEXT CONTINUATION TASKS
3771	VRDMG: VOCAL RESTORATION VIA DIFFUSION POSTERIOR SAMPLING WITH MULTIPLE GUIDANCE
3371	VT-REID: LEARNING DISCRIMINATIVE VISUAL-TEXT REPRESENTATION FOR POLYP RE-IDENTIFICATION
7958	Vulnerability of Face Age Verification to Replay Attacks
5306	WATER LEAK DETECTION VIA DOMAIN ADAPTATION
3576	WATERDIFF: PERCEPTUAL IMAGE WATERMARKS VIA DIFFUSION MODEL
2545	WAV2VEC-VC: VOICE CONVERSION VIA HIDDEN REPRESENTATIONS OF WAV2VEC 2.0
8500	WAVELET-DECOUPLING CONTRASTIVE ENHANCEMENT NETWORK FOR FINE-GRAINED SKELETON-BASED ACTION RECOGNITION



<b>Paper Number</b>	<b>Paper Title</b>
10071	Wavelet-Guided Acceleration of Text Inversion in Diffusion-Based Image Editing
3081	WAVELET-INSPIRED MULTISCALE GRAPH CONVOLUTIONAL RECURRENT NETWORK FOR TRAFFIC FORECASTING
2695	WAVER: WRITING-STYLE AGNOSTIC TEXT-VIDEO RETRIEVAL VIA DISTILLING VISION-LANGUAGE MODELS THROUGH OPEN-VOCABULARY KNOWLEDGE
2063	Weakly Semi-supervised Tool Detection in Minimally Invasive Surgery Videos
2758	Weakly Supervised Few-Shot Segmentation through Textual Prompt
8891	WEAKLY-SUPERVISED CROWD COUNTING WITH TOKEN ATTENTION AND FUSION: A SIMPLE AND EFFECTIVE BASELINE
3945	WFTNET: EXPLOITING GLOBAL AND LOCAL PERIODICITY IN LONG-TERM TIME SERIES FORECASTING
2938	WHAT DO NEURAL NETWORKS LISTEN TO? EXPLORING THE CRUCIAL BANDS IN SPEECH ENHANCEMENT USING SINC-CONVOLUTION
2043	WHAT DO SELF-SUPERVISED SPEECH AND SPEAKER MODELS LEARN? NEW FINDINGS FROM A CROSS MODEL LAYER-WISE ANALYSIS
9672	WHEN GREEN LEARNING MEETS FEDERATED LEARNING: TOWARD DISTRIBUTED LEARNING WITH LOW COMPLEXITY AND MODEL HETEROGENEITY
8965	WHEN TRAINING-FREE NAS MEETS VISION TRANSFORMERS: A NEURAL TANGENT KERNEL PERSPECTIVE
9880	WHICH IS THE BETTER TEACHER ACTION? A NEW RANKING MODEL AND DATASET
5374	WHISPER-BASED TRANSFER LEARNING FOR ALZHEIMER DISEASE CLASSIFICATION: LEVERAGING SPEECH SEGMENTS WITH FULL TRANSCRIPTS AS PROMPTS
11487	WHY DO ANGULAR MARGIN LOSSES WORK WELL FOR SEMI-SUPERVISED ANOMALOUS SOUND DETECTION?

<b>Paper Number</b>	<b>Paper Title</b>
1117	Widrow-Hoff LMS Adaline Demonstrator for Schools and Colleges
7777	Wi-Fi based Indoor Monitoring enhanced by Multimodal Fusion
8225	WiFiAct: Enhancing Human Sensing Through Environment Robust Preprocessing and Bayesian Self-Supervised Learning
7824	WiGig-based Joint Multi-Person Positioning and Respiration Sensing
1657	Window-based Convolutional Sparse Coding: Towards A Unified Framework
5080	X-CAUNET: CROSS-COLOR CHANNEL ATTENTION WITH UNDERWATER IMAGE-ENHANCING TRANSFORMER
11875	XIMALAYA ASDR SYSTEM FOR ICASSP 2024 IN-CAR MULTI-CHANNEL (ICMC) ASR CHALLENGE
2433	XMP: A Cross-Attention Multi-Scale Performer for File Fragment Classification
8596	YOLO-MED : MULTI-TASK INTERACTION NETWORK FOR BIOMEDICAL IMAGES
5836	ZE-FESG: A ZERO-SHOT FEATURE EXTRACTION METHOD BASED ON SEMANTIC GUIDANCE FOR NO-REFERENCE VIDEO QUALITY ASSESSMENT
3999	ZERO- AND FEW-SHOT SOUND EVENT LOCALIZATION AND DETECTION
1424	ZERO RESOURCE CODE-SWITCHED SPEECH BENCHMARK USING SPEECH UTTERANCE PAIRS FOR MULTIPLE SPOKEN LANGUAGES
2740	ZERO SHOT AUDIO TO AUDIO EMOTION TRANSFER WITH SPEAKER DISENTANGLEMENT
8138	Zero-Shot Co-salient object detection Framework
9568	Zero-shot Imitation Policy via Search in Demonstration Dataset
4270	ZERO-SHOT INTENT CLASSIFICATION USING A SEMANTIC SIMILARITY AWARE CONTRASTIVE LOSS AND LARGE LANGUAGE MODEL

Paper Number	Paper Title
3262	ZERO-SHOT OBJECT DETECTION WITH PARTITIONED CONTRASTIVE FEATURE ALIGNMENT
8834	ZIGZAG ATTENTION: A STRUCTURAL AWARE MODULE FOR LANE DETECTION
6709	ZIV-ZAKAI BOUND FOR DOA ESTIMATION WITH GAIN-PHASE ERROR

[Contact](#) | [Accessibility](#) | [Nondiscrimination Policy](#) | [IEEE Ethics Reporting](#) | [IEEE Privacy Policy](#) | [Terms](#) | [Signal Processing Society](#)

©2024 [IEEE](#) – All rights reserved.

Use of this website signifies your agreement to the [IEEE Terms and Conditions](#).

Last updated Last updated 13 December 2023.

Support: [info@2024.ieeeicassp.org](mailto:info@2024.ieeeicassp.org) Host: <https://cmsworldwide.com/>