

Dataset

UGC contents

309 Processed Video Sources (PVS)

16 sources (SRC)

720p, vertical, 30 fps
5 to 20 second clips

3 Constant Rate Factors (CRF)

6 Pre-processing (combination of)

Artifact Removal (AR)

mctd-ar-sharp-low,
mctd-ar-sharp-med

Noise Reduction (MCTD)

mctd-ar-sharp-high
mctd-sharp-low

Sharpness (3 levels)

ar-sharp-low
ar-mctd-sharp-low

4 codecs

H264 vs H265

With and without ROI encoding

Selection

Bitrate-distortion (BD) behavior-based

Multi-codec

Multi-quality metric

Multi-scale

Multi-pre-processing

Priority to contents with special effects, filters

Presence of contents with Chinese characters

Ratings

ACR scores

20 scores per PVS per population

Two sets of population

Chinese speakers

Non Chinese speakers

800 observers overall

5-minutes online test

Benchmark

29 Metrics

VMAF crf1_VMAF NIMA VSFA
PSNR* crf1_PSNR* NIMA_mp BRISQUE
SSIM* crf1_SSIM* MLSP NIQE
RAPIQUE VIDEVAL ILNIQE VIIDEO
VQSCORE

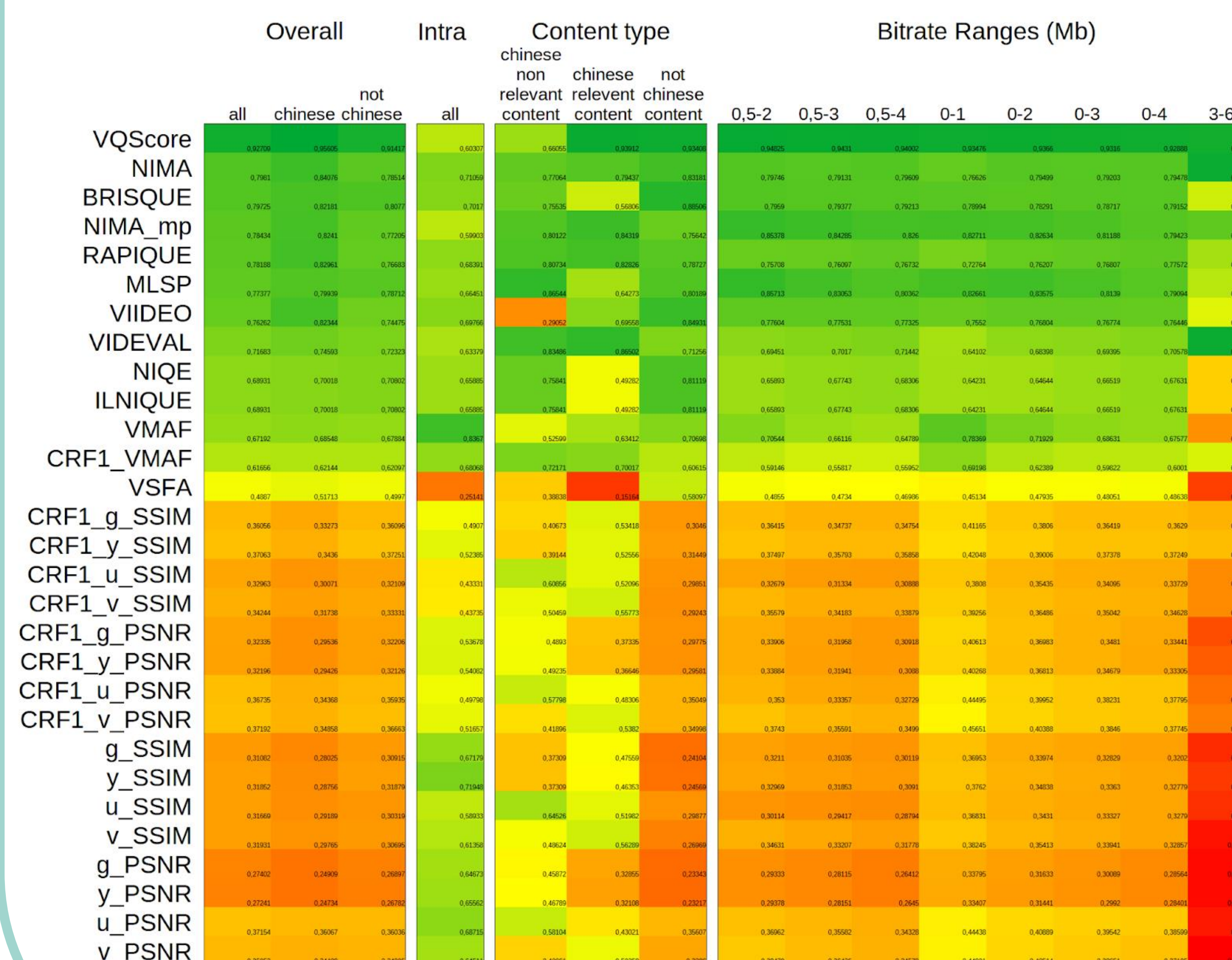
* in 4 channels: Y, U, V, and G (global).

5 Figures of Merit

PLCC SROCC KROCC RMSE C0

Conditions

Intra VS Inter pairs
Content type Population
Bitrate Ranges Overall



Metrics Combination

Machine Learning Algorithms

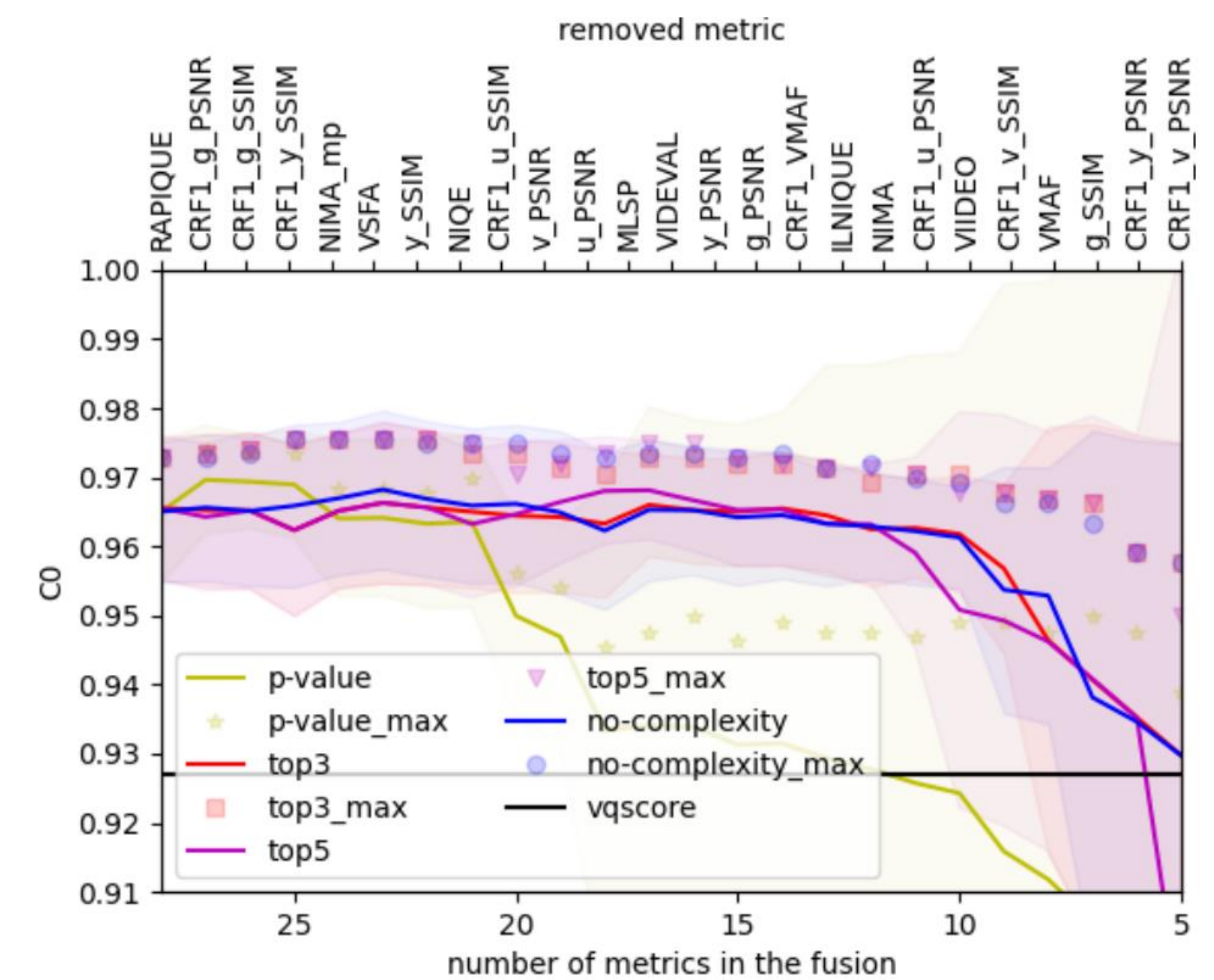
SVM SVR Random Forest Adaboost

Dimension reduction strategies

No-complexity p-value Top 3 Top 5

Backward Feature Elimination

From 29 to 5 metrics



Final Fusion Metric

BRISQUE {V, U, G} SSIM SVR C0
VQSCORE crf1_{Y, V} PSNR with No-complexity

	Metrics fusion	VQScore	VMAF
PLCC	0.9339	0.8632	0.3496
SRCC	0.9297	0.8597	0.3734
KROCC	0.7700	0.6672	0.2544
C0	0.9656	0.9271	0.6719