

WHAT IS CLAIMED IS:

1. A computer-implemented method, comprising:
for each genre included in a plurality of genres, generating a partition that
subdivides a plurality of metric values associated with baseline encoded
5 video content based on the genre;
performing one or more cross-validation operations across the partitions to
determine optimized hyperparameters for a model;
performing one or more training operations on the model based on the
optimized hyperparameters and the plurality of metric values to
10 generate a trained model; and
applying the trained model to target encoded video content to compute a
perceptual video quality metric value.
2. The computer-implemented method of claim 1, wherein performing the one or
15 more training operations on the model comprises applying a random forest algorithm
to one or more objective metric values included in the plurality of metric values.
3. The computer-implemented method of claim 1, wherein the plurality of metric
values includes values for at least one of Blind/Referenceless Image Spatial Quality
20 Evaluator, temporal information, and visual information fidelity.
4. The computer-implemented method of claim 1, wherein the plurality of genres
includes at least one of an animation genre, an action genre, and a heavy-noise
25 genre.
5. The computer-implemented method of claim 1, wherein the plurality of metric
values includes at least one human-observed score of visual quality for reconstructed
video content that is derived from the baseline encoded video content.
- 30 6. The computer-implemented method of claim 1, wherein for a given genre,
generating the partition comprises:
determining that a first genre included in the plurality of genres and associated
with a first metric value included in the plurality of metric values is equal
to the given genre;

adding the first metric value to testing metric values included in the partition;
determining that a second genre included in the plurality of genres and
associated with a second metric value included in the plurality of metric
values is not equal to the given genre; and
5 adding the second metric value to training metric values included in the
partition.

7. The computer-implemented method of claim 1, wherein performing one or
more cross-validation operations comprises:
10 selecting first hyperparameters for the model;
computing a first correlation based on the first hyperparameters and the
partitions;
determining that the first correlation is closer than a second correlation
associated with second hyperparameters for the model; and
15 setting the optimized hyperparameters equal to the first hyperparameters.

8. The computer-implemented method of claim 7, wherein computing the first
correlation comprises:
for each partition, pre-training the model based on the first hyperparameters
20 and one or more training metric values included in the partition to generate
a partition-specific model;
for each partition, applying the partition-specific model to one or more testing
metric values included in the partition to compute predicted perceptual
video quality scores; and
25 performing one or more correlation operations between the predicted
perceptual visual quality scores and one or more subjective metric values
included in the plurality of metric values.

9. A computer-readable storage medium including instructions that, when
30 executed by a processor, cause the processor to perform the steps of:
for each genre included in a plurality of genres, generating a partition that
subdivides baseline data that describes baseline encoded video content
based on the genre;
generating a plurality of hyperparameter combinations for a model;

- performing one or more cross-validation operations across the partitions to
select optimized hyperparameters from the plurality of hyperparameter
combinations;
- performing one or more training operations on the model based on the
5 optimized hyperparameters and the baseline data to generate a trained
model; and
- applying the trained model to target encoded video content to compute a
perceptual video quality metric value.
- 10 10. The computer-readable storage medium of claim 9, wherein the baseline data
includes values for at least one of Blind/Referenceless Image Spatial Quality
Evaluator, temporal information, and visual information fidelity.
11. The computer-readable storage medium of claim 9, wherein the baseline data
15 includes at least one human-observed score of visual quality for reconstructed video
content that is derived from the baseline encoded video content.
12. The computer-readable storage medium of claim 9, wherein for a first genre,
generating the partition comprises:
- 20 determining that a second genre included in the plurality of genres and
associated with a first data included in the baseline data is equal to the
first genre;
- adding the first data to testing metric values included in the partition;
- determining that a third genre included in the plurality of genres and associated
25 with a second data included in the baseline data is not equal to the first
genre; and
- adding the second data to training data included in the partition.
13. The computer-readable storage medium of claim 9, wherein performing one or
30 more cross-validation operations comprises:
- selecting a first hyperparameter combination included in the plurality of
hyperparameter combinations;
- computing a first correlation based on the first hyperparameter combination
and the partitions;

determining that the first correlation is closer than a second correlation associated with a second hyperparameter combination included in the plurality of hyperparameter combinations; and
setting the optimized hyperparameters equal to the first hyperparameter combination.

14. The computer-readable storage medium of claim 13, wherein the first correlation comprises at least one of a root mean squared error, a Pearson correlation coefficient, and a Spearman's rank order correlation coefficient.

10

15. The computer-readable storage medium of claim 9, wherein performing one or more cross-validation operations comprises:

for each partition, pre-training the model based on a first hyperparameter combination included in the plurality of hyperparameter combinations and one or more training metric values included in the partition to generate a partition-specific model;

15

for each partition, applying the partition-specific model to one or more testing metric values included in the partition to compute predicted perceptual video quality scores; and

20

performing one or more correlation operations between the predicted perceptual visual quality scores and one or more subjective metric values included in the baseline data.

16. The computer-readable storage medium of claim 15, wherein pre-training the model comprises applying a random forest algorithm that implements the first hyperparameter combination to one or more objective metric values included in the one or more training metric values.

25

17. A system comprising:
a memory storing instructions associated with a perceptual visual quality modeling engine; and
a processor that is coupled to the memory and, when executing the instructions, is configured to:
for each genre included in a plurality of genres, generate a partition that

30

subdivides a plurality of metric values associated with baseline encoded video content based on the genre;
for each partition, perform one or more pre-training operations to compute predicted perceptual quality scores;
5 determine optimized hyperparameters for a model based on the predicted perceptual quality scores associated with the partitions;
perform one or more training operations on the model based on the optimized hyperparameters and the plurality of metric values to generate a trained model; and
10 apply the trained model to target encoded video content to compute an output perceptual quality score.

18. The system of claim 17, wherein the plurality of metric values includes at least one human-observed score of visual quality for reconstructed video content that is
15 derived from the baseline encoded video content.

19. The system of claim 17, wherein the plurality of genres includes at least one of an animation genre, an action genre, and a heavy-noise genre.

20. The system of claim 17, wherein the target encoded video data includes at least one of compressed data and scaled data.