Fig. 1: Two sets of example blocks cropped from the reconstructed frames generated by the anchor HM 16.20 (QP=37), six state-of-the-art network architectures and the proposed CVEGAN for CNN-based PP. The bit consumption in each example set is identical for all tested versions. Row 1 and 2 correspond to the 170th frame of the *PartyScene* sequence and Row 3 and 4 correspond to the 104th frame of the *CatRobot1* sequence. It can be observed that the output of CVEGAN exhibits improved perceptual quality compared to the anchor HEVC HM 16.20 and other compared networks, with fewer blocking artefacts, more textural detail and higher contrast.
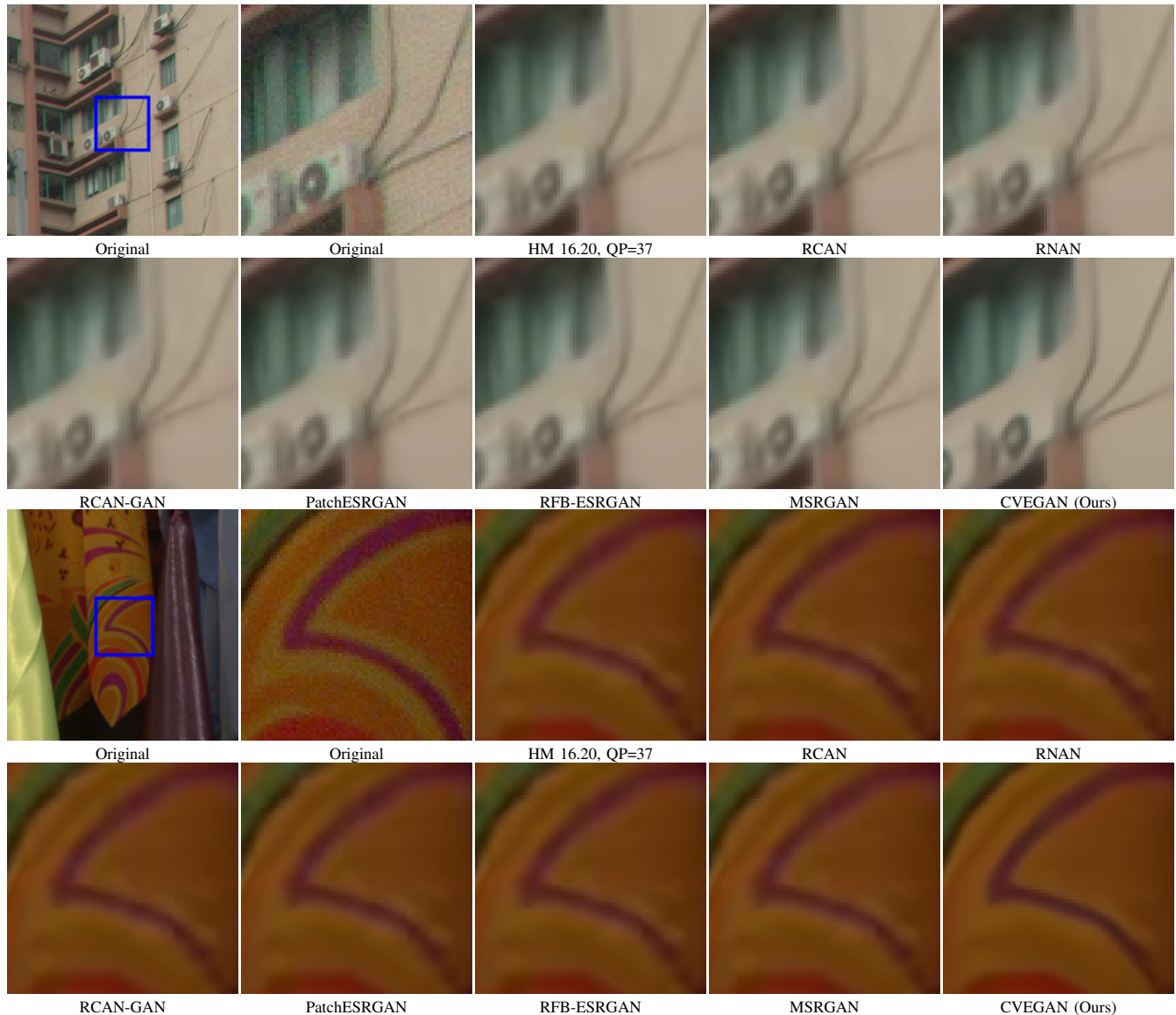
Fig. 2: Two sets of example blocks cropped from the reconstructed frames generated by the anchor HM 16.20 (QP=37), six state-of-the-art network architectures and the proposed CVEGAN for CNN-based PP. The bit consumption in each example set is identical for all tested versions. Row 1 and 2 correspond to the 250th frame of the *DaylightRoad2* sequence, Row 3 and 4 correspond to the 216th frame of the *CatRobot1* sequence. It can be observed that the output of CVEGAN exhibits improved perceptual quality compared to the anchor HEVC HM 16.20 and other compared networks, with fewer blocking artefacts, more textural detail and higher contrast.
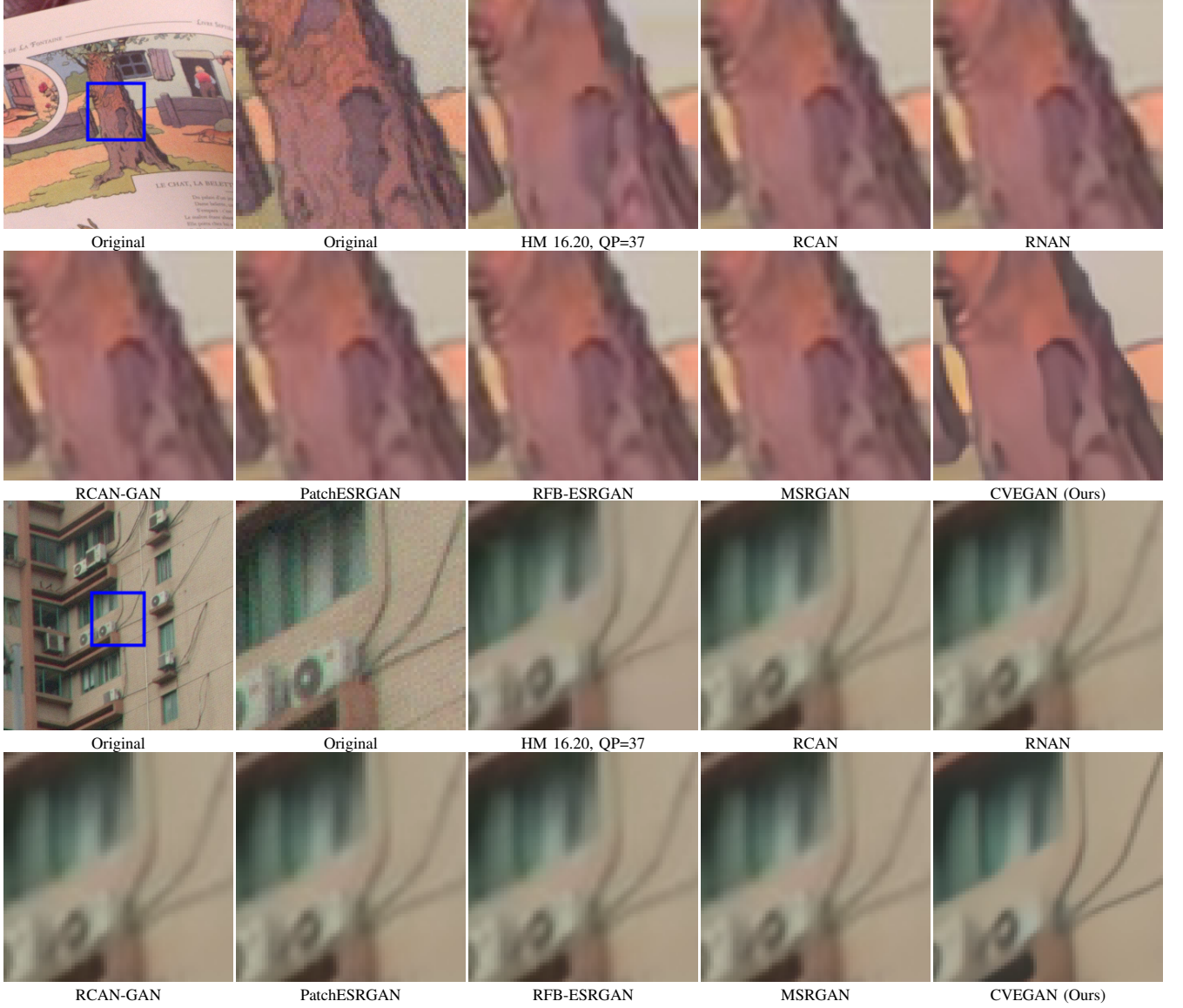
Fig. 3: Two sets of example blocks cropped from the reconstructed frames generated by the anchor HM 16.20 (QP=37), six state-of-the-art network architectures and the proposed CVEGAN for CNN-based SRA. The bit consumption in each example set is similar for all tested versions. Row 1 and 2 correspond to the 104th frame of the *CatRobot1* sequence, Row 3 and 4 correspond to the 250th frame of the *DaylightRoad2* sequence. It can be observed that the output of CVEGAN exhibits improved perceptual quality compared to the anchor HEVC HM 16.20 and other networks, with fewer blocking artefacts, more textural detail and higher contrast.
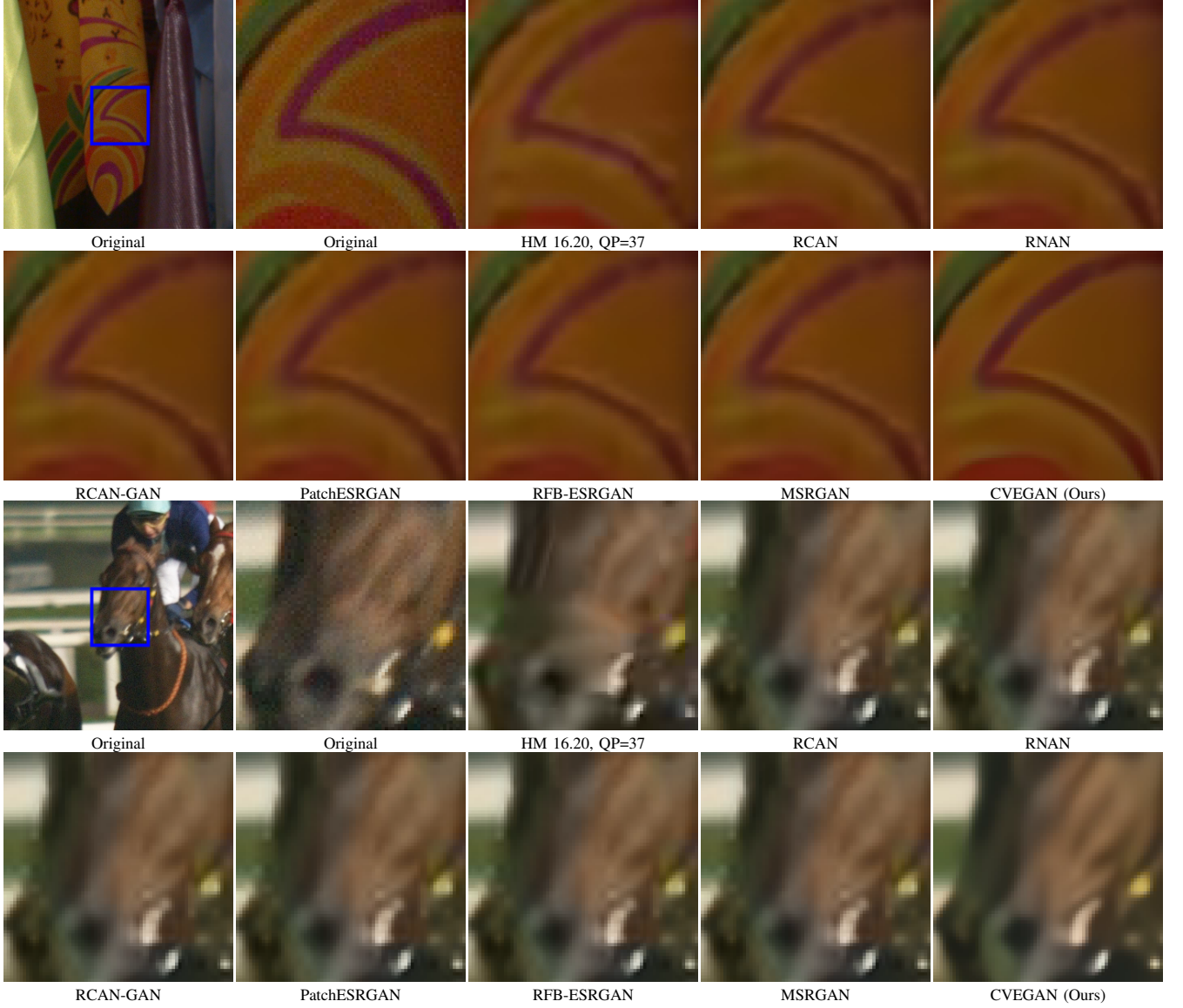
Fig. 4: Two sets of example blocks cropped from the reconstructed frames generated by the anchor HM 16.20 (QP=37), six state-of-the-art network architectures and the proposed CVEGAN for CNN-based SRA. The bit consumption in each example set is similar for all tested versions. Row 1 and 2 correspond to the 216th frame of the *CatRobot1* sequence and Row 3 and 4 correspond to the 161st frame of the *RaceNight* sequence. It can be observed that the output of CVEGAN exhibits improved perceptual quality compared to the anchor HEVC HM 16.20 and other networks, with fewer blocking artefacts, more textural detail and higher contrast.