**TableS9 : Average accuracies obtained through video-unit-based predictions grouped in terms of dataset and network types**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Efficient  Net B0 | Mobile  Net V2 | Xception | Ave | MaxAve | MinAve | Gap |
| Set X | 0.8493 | 0.8373 | 0.8318 | 0.8395 | 0.8914 | 0.8395 | 0.0519 |
| Set Y | 0.8839 | 0.8871 | 0.9031 | 0.8914 |
| Ave | 0.8666 | 0.8622 | 0.8675 |  |  |  |  |
| MaxAve | 0.8675 | | |  |  |  |  |
| MinAve | 0.8622 | | |  |  |  |  |
| Gap | 0.0052 | | |  |  |  |  |