**Results from testing under training setting TS3** (models trained on the **whole**  4comb dataset and then tested on another dataset – **ccdb**)

* testing nod, shake, and tilt models independently as well as a final fused 4-class (none/nod/shake/tilt) HGD model
* results from models trained with and without data augmentation (A1)
* testing using non-smoothed and smoothed predictions
* number of features is fixed to 12
* window size is fixed to 32 frames
* uses the best model architectures found in TS2 during validation
* the results were generated using the scripts deep-virtual-rapport-agent/head\_gesture\_detector/evaluate\_final\_4comb\_hgd.ipynb and deep-virtual-rapport-agent/head\_gesture\_detector/evaluate\_fused\_final\_4comb\_hgd.ipynb

**Independent nod, shake, and tilt models**

**NO AUGMENTATION**

**NON-SMOOTHED**

**NOD**

Loading model:

final\_4comb\_nod\_32ws\_12f\_16u.hdf5 #params: 1409

Testing on ccdb

[last] bacc: 0.8024

[last] f1: 0.2236

[last] precision: 0.1278

[last] recall: 0.8950

[majority] bacc: 0.8038

[majority] f1: 0.2240

[majority] precision: 0.1279

[majority] recall: 0.8984

Total time taken: 11.967236757278442 s

Saved to ./test\_results\_final\_4comb\_nonS/test\_results\_final\_4comb\_nonS\_nod\_32ws\_12f\_16u.pkl

**SHAKE**

Loading model:

final\_4comb\_shake\_32ws\_12f\_8u.hdf5 #params: 513

Testing on ccdb

[last] bacc: 0.8263

[last] f1: 0.1068

[last] precision: 0.0570

[last] recall: 0.8435

[majority] bacc: 0.8388

[majority] f1: 0.1173

[majority] precision: 0.0630

[majority] recall: 0.8508

Total time taken: 13.208979606628418 s

Saved to ./test\_results\_final\_4comb\_nonS/test\_results\_final\_4comb\_nonS\_shake\_32ws\_12f\_8u.pkl

**TILT**

Loading model:

final\_4comb\_tilt\_32ws\_12f\_16u.hdf5 #params: 1409

Testing on ccdb

[last] bacc: 0.8013

[last] f1: 0.0710

[last] precision: 0.0373

[last] recall: 0.7502

[majority] bacc: 0.7950

[majority] f1: 0.0707

[majority] precision: 0.0371

[majority] recall: 0.7354

Total time taken: 12.079846620559692 s

Saved to ./test\_results\_final\_4comb\_nonS/test\_results\_final\_4comb\_nonS\_tilt\_32ws\_12f\_16u.pkl

**SMOOTHED**

**NOD**

Loading model:

final\_4comb\_nod\_32ws\_12f\_16u.hdf5 #params: 1409

Testing on ccdb

[last] bacc: 0.8068

[last] f1: 0.2274

[last] precision: 0.1302

[last] recall: 0.8991

[majority] bacc: 0.8076

[majority] f1: 0.2276

[majority] precision: 0.1302

[majority] recall: 0.9011

Total time taken: 11.050076723098755 s

Saved to ./test\_results\_final\_4comb\_S/test\_results\_final\_4comb\_S\_nod\_32ws\_12f\_16u.pkl

**SHAKE**

Loading model:

final\_4comb\_shake\_32ws\_12f\_8u.hdf5 #params: 513

Testing on ccdb

[last] bacc: 0.8300

[last] f1: 0.1092

[last] precision: 0.0584

[last] recall: 0.8471

[majority] bacc: 0.8458

[majority] f1: 0.1214

[majority] precision: 0.0653

[majority] recall: 0.8601

Total time taken: 11.328006267547607 s

Saved to ./test\_results\_final\_4comb\_S/test\_results\_final\_4comb\_S\_shake\_32ws\_12f\_8u.pkl

**TILT**

Loading model:

final\_4comb\_tilt\_32ws\_12f\_16u.hdf5 #params: 1409

Testing on ccdb

[last] bacc: 0.8053

[last] f1: 0.0740

[last] precision: 0.0389

[last] recall: 0.7521

[majority] bacc: 0.7985

[majority] f1: 0.0740

[majority] precision: 0.0389

[majority] recall: 0.7354

Total time taken: 12.301400899887085 s

Saved to ./test\_results\_final\_4comb\_S/test\_results\_final\_4comb\_S\_tilt\_32ws\_12f\_16u.pkl

WITH AUGMENTATION A1

**NON-SMOOTHED**

**NOD**

Loading model:

final\_4comb\_nod\_32ws\_12f\_A1\_16u.hdf5 #params: 1409

Testing on ccdb

[last] bacc: 0.7926

[last] f1: 0.2169

[last] precision: 0.1236

[last] recall: 0.8823

[majority] bacc: 0.7938

[majority] f1: 0.2176

[majority] precision: 0.1241

[majority] recall: 0.8842

Total time taken: 13.328564405441284 s

Saved to ./test\_results\_final\_4comb\_nonS/test\_results\_final\_4comb\_nonS\_nod\_32ws\_12f\_A1\_16u.pkl

**SHAKE**

Loading model:

final\_4comb\_shake\_32ws\_12f\_A1\_8u.hdf5 #params: 513

Testing on ccdb

[last] bacc: 0.8228

[last] f1: 0.1030

[last] precision: 0.0548

[last] recall: 0.8450

[majority] bacc: 0.8349

[majority] f1: 0.1109

[majority] precision: 0.0593

[majority] recall: 0.8554

Total time taken: 13.143837213516235 s

Saved to ./test\_results\_final\_4comb\_nonS/test\_results\_final\_4comb\_nonS\_shake\_32ws\_12f\_A1\_8u.pkl

**TILT**

Loading model:

final\_4comb\_tilt\_32ws\_12f\_A1\_16u.hdf5 #params: 1409

Testing on ccdb

[last] bacc: 0.7639

[last] f1: 0.0640

[last] precision: 0.0336

[last] recall: 0.6760

[majority] bacc: 0.7759

[majority] f1: 0.0670

[majority] precision: 0.0352

[majority] recall: 0.6973

Total time taken: 13.736355543136597 s

Saved to ./test\_results\_final\_4comb\_nonS/test\_results\_final\_4comb\_nonS\_tilt\_32ws\_12f\_A1\_16u.pkl

**SMOOTHED**

**NOD**

Loading model:

final\_4comb\_nod\_32ws\_12f\_A1\_16u.hdf5 #params: 1409

Testing on ccdb

[last] bacc: 0.7948

[last] f1: 0.2187

[last] precision: 0.1248

[last] recall: 0.8842

[majority] bacc: 0.7984

[majority] f1: 0.2210

[majority] precision: 0.1262

[majority] recall: 0.8891

Total time taken: 12.261885643005371 s

Saved to ./test\_results\_final\_4comb\_S/test\_results\_final\_4comb\_S\_nod\_32ws\_12f\_A1\_16u.pkl

**SHAKE**

Loading model:

final\_4comb\_shake\_32ws\_12f\_A1\_8u.hdf5 #params: 513

Testing on ccdb

[last] bacc: 0.8224

[last] f1: 0.1040

[last] precision: 0.0555

[last] recall: 0.8409

[majority] bacc: 0.8360

[majority] f1: 0.1128

[majority] precision: 0.0604

[majority] recall: 0.8539

Total time taken: 11.787539005279541 s

Saved to ./test\_results\_final\_4comb\_S/test\_results\_final\_4comb\_S\_shake\_32ws\_12f\_A1\_8u.pkl

**TILT**

Loading model:

final\_4comb\_tilt\_32ws\_12f\_A1\_16u.hdf5 #params: 1409

Testing on ccdb

[last] bacc: 0.7655

[last] f1: 0.0665

[last] precision: 0.0350

[last] recall: 0.6722

[majority] bacc: 0.7832

[majority] f1: 0.0713

[majority] precision: 0.0376

[majority] recall: 0.7038

Total time taken: 13.771583557128906 s

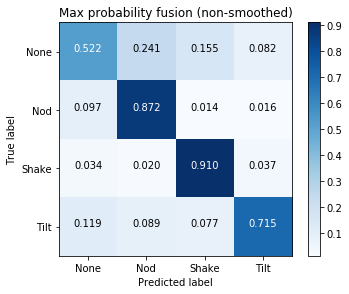
Saved to ./test\_results\_final\_4comb\_S/test\_results\_final\_4comb\_S\_tilt\_32ws\_12f\_A1\_16u.pkl

**Final fused 4-class (none/nod/shake/tilt) HGD model**

**[currently implemented] Skipping frames with multiple head gesture annotations in ground truths**

**Max probabiliy fusion & non-smoothed**

**Normalized confusion matrix**



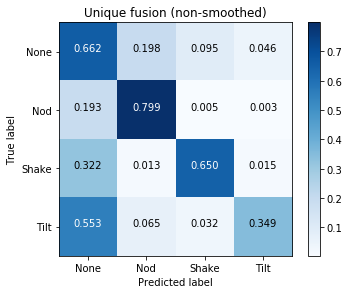
**Recall: 0.5432399341373827**

**Precision: 0.9360818445593223**

**F1: 0.6546629948009215**

**Unique fusion & non-smoothed**

**Normalized confusion matrix**



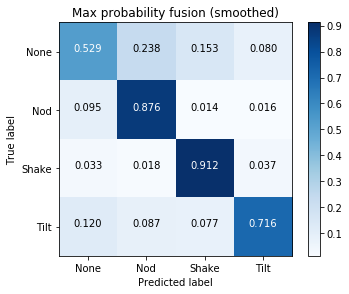
**Recall: 0.665894985715693**

**Precision: 0.924447111347626**

**F1: 0.7541276923248904**

**Max probabiliy fusion & smoothed**

**Normalized confusion matrix**



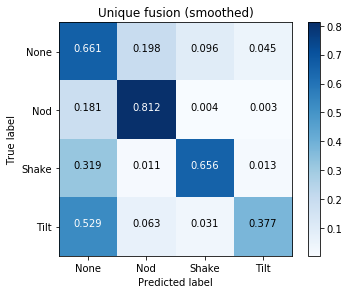
**Recall: 0.5496150977384353**

**Precision: 0.9365270836099255**

**F1: 0.6601689672542163**

**Unique fusion & smoothed**

**Normalized confusion matrix**



**Recall: 0.666584572948478**

**Precision: 0.9255477555065331**

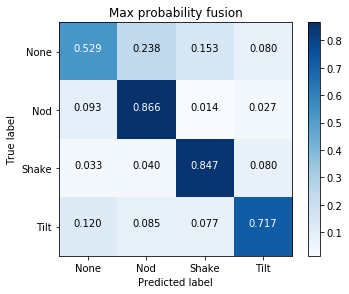
**F1: 0.7545617314797356**

Compare with the work “Recognizing Visual Signatures of Spontaneous Head Gestures” <https://ieeexplore.ieee.org/document/8354154>

**[earlier attempts below] with individual binary predictions smoothed**

**GROUND TRUTHS FUSED imposing priorities nod > shake > tilt if a frame is annotated with 2 or more head gestures at a time (e.g. nod and shake at the same time).**

Normalized confusion matrix

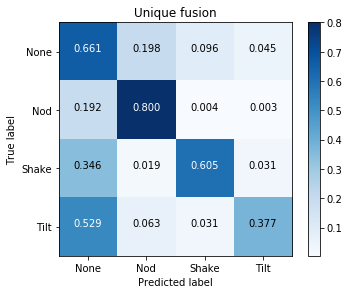


Recall: 0.5491147273978181

Precision: 0.9345386023224717

F1: 0.6591183099783696

Normalized confusion matrix



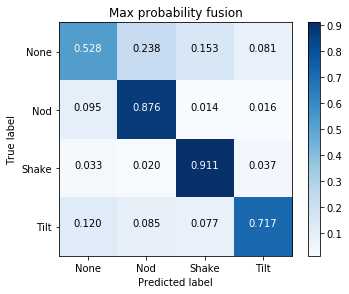
Recall: 0.6655200011232643

Precision: 0.9216536308724117

F1: 0.7526325247863189

**GROUND TRUTHS FUSED using the none label when a frame is annotated with 2 or more head gestures at a time (e.g. nod and shake at the same time).**

Normalized confusion matrix

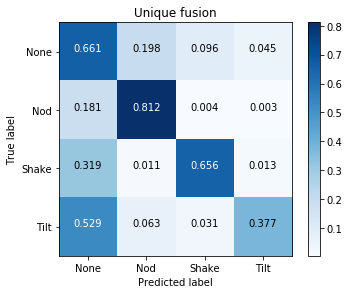


Recall: 0.5484056668679181

Precision: 0.9366348424534289

F1: 0.6592405753160268

Normalized confusion matrix



Recall: 0.6664958369020373

Precision: 0.9257007460156367

F1: 0.7545808387134513