

ЕМ-алгоритм для разделения смесей нормальных распределений

1 Введение

В данной работе был проведен анализ работы ЕМ-алгоритма, разделяющего смеси нормальных распределений, на примере задачи вычитания фона. ЕМ-алгоритм был реализован на языке python. Эксперименты были проведены на датасете pedestrians. В качестве мер качества использовались FN, FP, ROC AUC, а также визуальная оценка качества.

2 Постановка задачи

Рассматривалась задача классификации пикселей изображения на два класса: относящиеся к переднему плану и относящиеся к фону. Были рассмотрены три модели: модель с одномерной гауссианой, модель с трехмерной гауссианой и модель со смесью трехмерных гауссиан.

Первая модель предполагает, что яркость каждого пикселя фона имеет нормальное распределение с неменяющимися со временем средним и дисперсией, и что яркости пикселей, не относящихся к фону сильно отклоняются от этих средних. Идея второй модели такая же, как и у первой, но здесь яркость рассматривается отдельно по каждому из каналов. Третья модель не требует унимодальности распределений.

3 Эксперименты

3.1 Проверка работы алгоритма

Перед запуском алгоритма на реальных данных проверим, как он разделяет смеси. На рисунке 1 приведены две смеси. Цветами показаны результаты разделения после работы алгоритма. Видно, что алгоритм хорошо разделил разные смеси. На рисунке 2 приведена одна и та же смесь и результаты четырех запусков алгоритма для нее. Видно, что ответ неустойчивый: при разных запусках у одной и той же компоненты получаются разные лейблы, а на некоторых запусках некоторые компоненты смеси вообще не разделились. Это объясняется тем, что ЕМ-алгоритм может попадать в локальные максимумы правдоподобия.

3.2 Одномерная гауссиана

Чтобы сделать предсказание для пикселя о том, относится он к фону или к переднему плану, рассматривалось отклонение его яркости, вычисленной по формуле $0.2126r + 0.7152g + 0.0722b$, от предсказанного математического ожидания. Это отклонение сравнивалось с числом $k * \sigma$ (σ^2 — предсказанная дисперсия), и если отклонение оказывалось больше, то пиксель относился

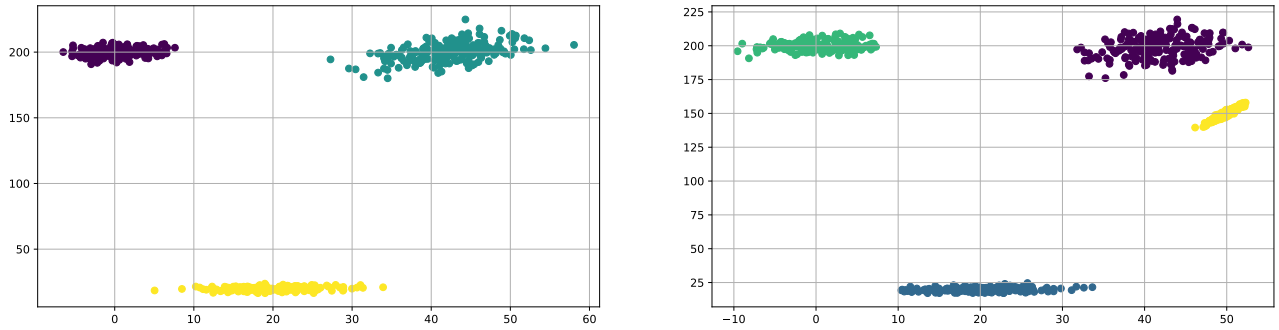


Рис. 1: Разделенные смеси

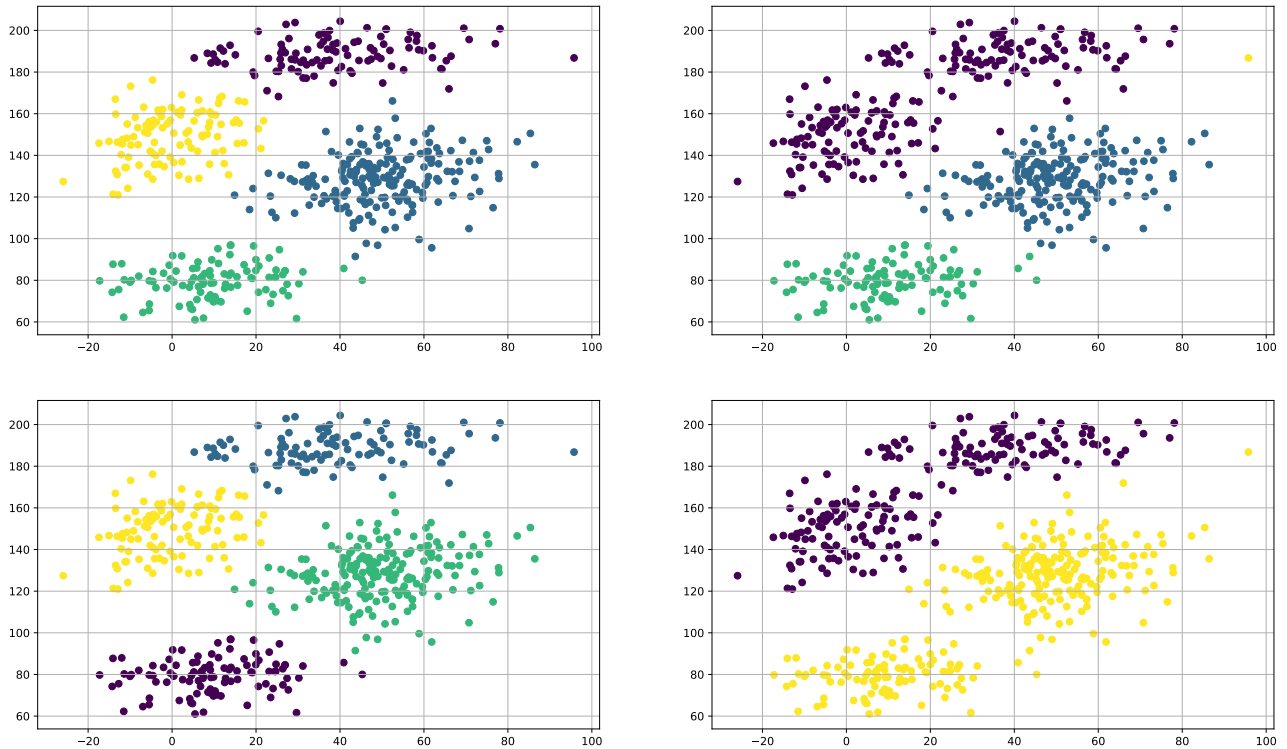


Рис. 2: Плохо разделенные смеси

к переднему плану, иначе к фону. На рисунке 3 приведены графики зависимости числа ошибок 2 и 1 рода от номера кадра. Разные кривые соответствуют разным значениям коэффициента k . Чем больше k , тем меньше пикселей будет отнесено к переднему плану (к отрицательному классу) и тем больше — к фону (к положительному классу). Соответственно тем больше FP и тем меньше FN, что и видно на графиках.

Также заметно, что примерно первую половину времени FN мало, а FP велико, и наоборот в оставшееся время. Это может объясняться тем, что за время съемки поменялось освещение и многие пиксели, относящиеся к фону, стали сильно отклоняться от своих старых предсказанных средних и были неверно классифицированы не как фоновые.

На рисунке 4 показаны лучший и худший по значению ROC-AUC кадры. На плохом кадре видно, что много ошибок появляется на границах и тенях объектов. Это объясняется тем, что пиксели, относящиеся к освещенному участку и, соответственно, полученные из одного распре-

деления, из-за смещения света начинают относиться к участку тени, и в результате алгоритм их выделяет как изменяющиеся объекты. Эту проблему должна помочь решить смесь гауссиан, так как она допускает наличие разных участков фона.

На рисунке 5 приведены графики ROC-кривых для разных кадров.

3.3 Трехмерная гауссиана

В этой модели три цветовых канала рассматривались отдельно. Решение о классификации пикселей принималось по значению плотности восстановленного распределения, которое сравнивалось с некоторым порогом. На рисунке 6 приведены графики зависимости числа ошибок 2 и 1 рода от номера кадра. Разные кривые соответствуют разным значениям порога плотности. Большие значения порога соответствуют тому, что меньше пикселей будет отнесено к переднему плану и больше пикселей — к фону. Следовательно, FP при этом увеличивается, а FN уменьшается, что и видно на графиках.

На рисунке 7 показаны лучший и худший по значению ROC-AUC кадры. Видно, что ошибок на границах на худшей картинке стало значительно меньше по сравнению с первой моделью и минимальный ROC-AUC стал значительно больше (0.38 вместо 0.12).

На рисунке 8 приведены графики ROC-кривых для разных кадров.

3.4 Смесь гауссиан

Здесь, как и в предыдущей модели, решение о классификации пикселей принималось по значению плотности, но уже смеси трех гауссиан. На рисунке 9 приведены графики зависимости числа ошибок 2 и 1 рода от номера кадра. Наблюдаются те же эффекты, что и в предыдущем эксперименте.

На рисунке 10 показаны лучший и худший по значению ROC-AUC кадры. Как и ожидалось, эта модель хорошо исправила ошибки первой модели. Хотя трехмерная гауссиана показала примерно такие же результаты. На рисунке 11 приведены графики ROC-кривых для разных кадров.

В прилагающемся к отчету html-файле приведен видеоряд с результатами работы последней модели как лучшей из рассмотренных.

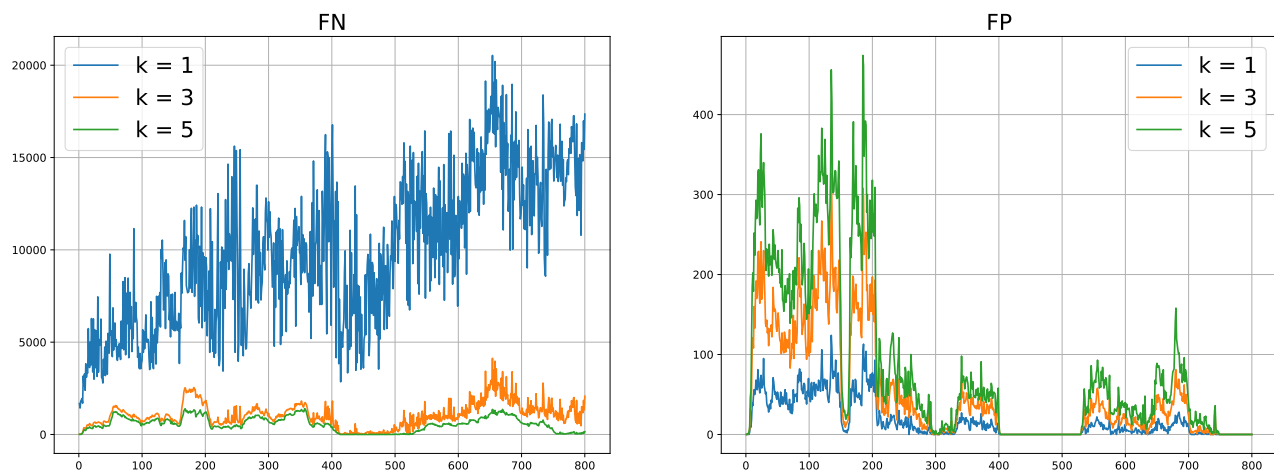


Рис. 3: Графики FN и FP для разных значений k

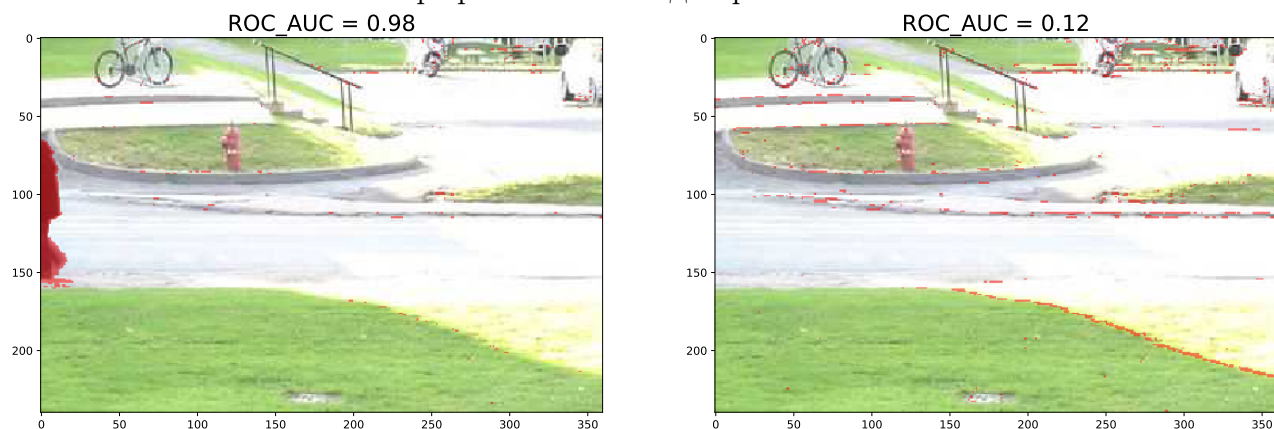


Рис. 4: Лучший и худший кадры в первой модели
ROC-кривые для разных моментов времени

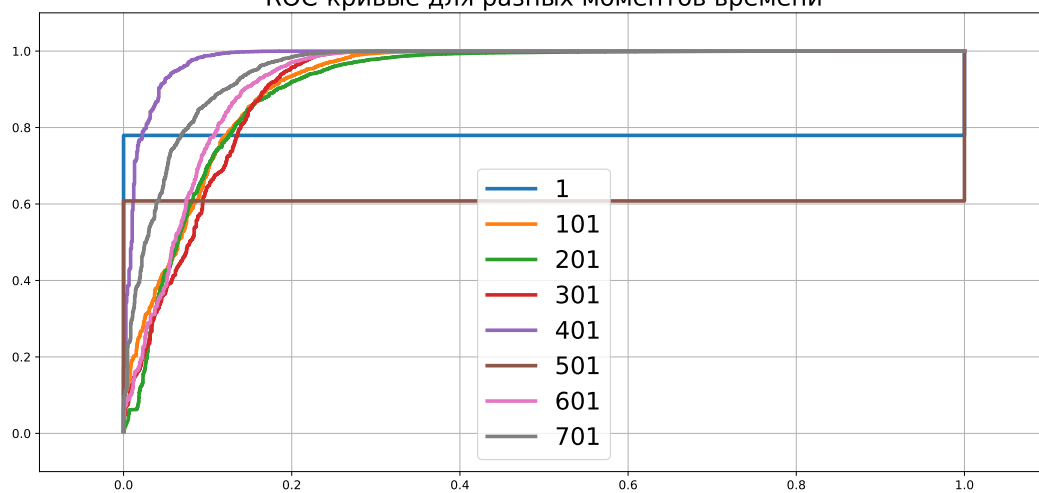


Рис. 5: Графики ROC-кривых для одномерной гауссианы

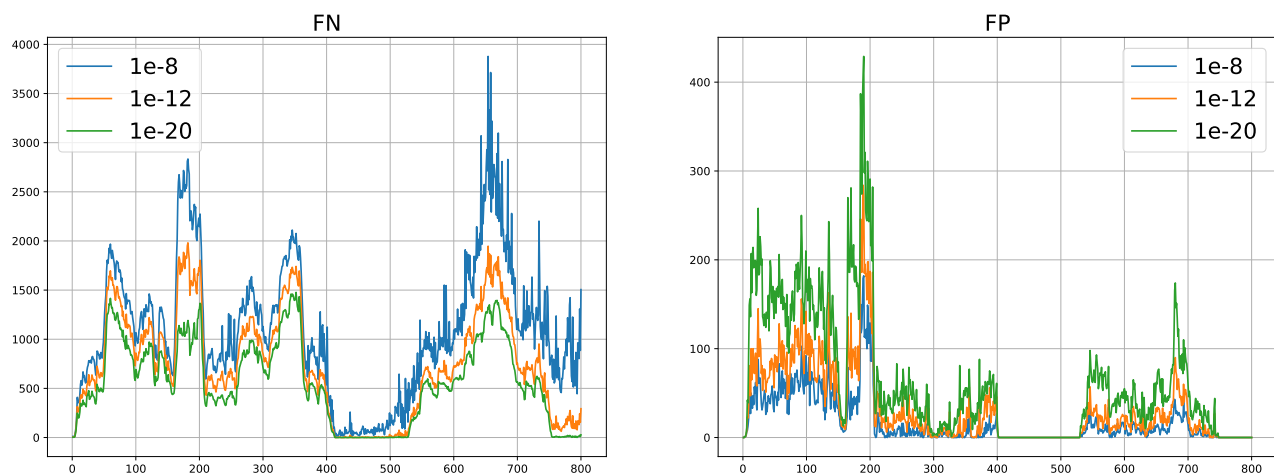


Рис. 6: Графики FN и FP для разных значений k

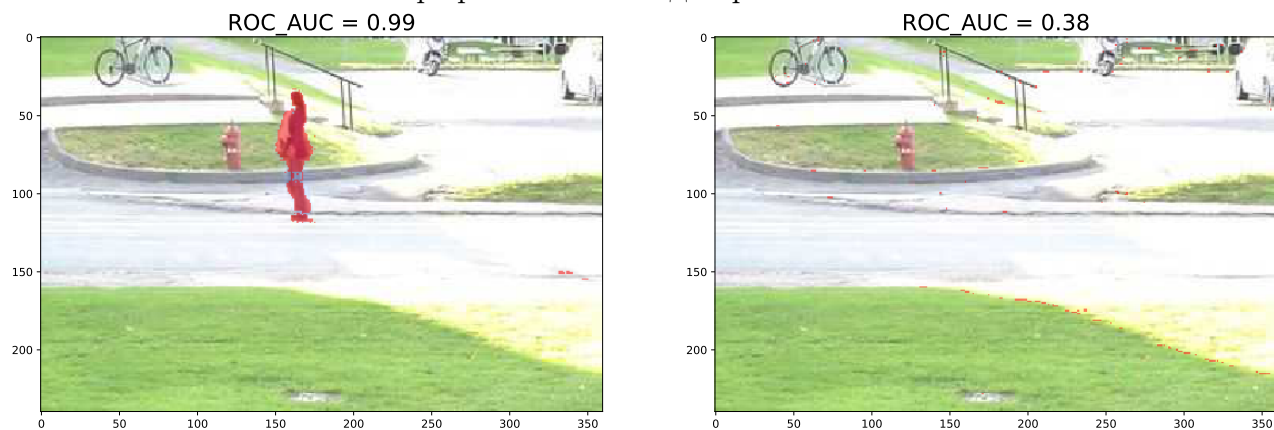


Рис. 7: Лучший и худший кадры во второй модели
ROC-кривые для разных моментов времени

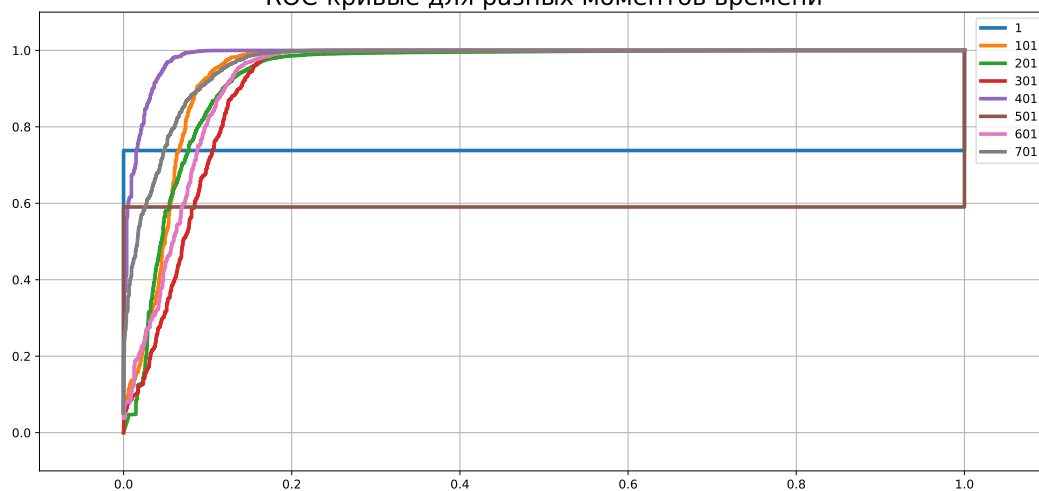


Рис. 8: Графики ROC-кривых для трехмерной гауссианы

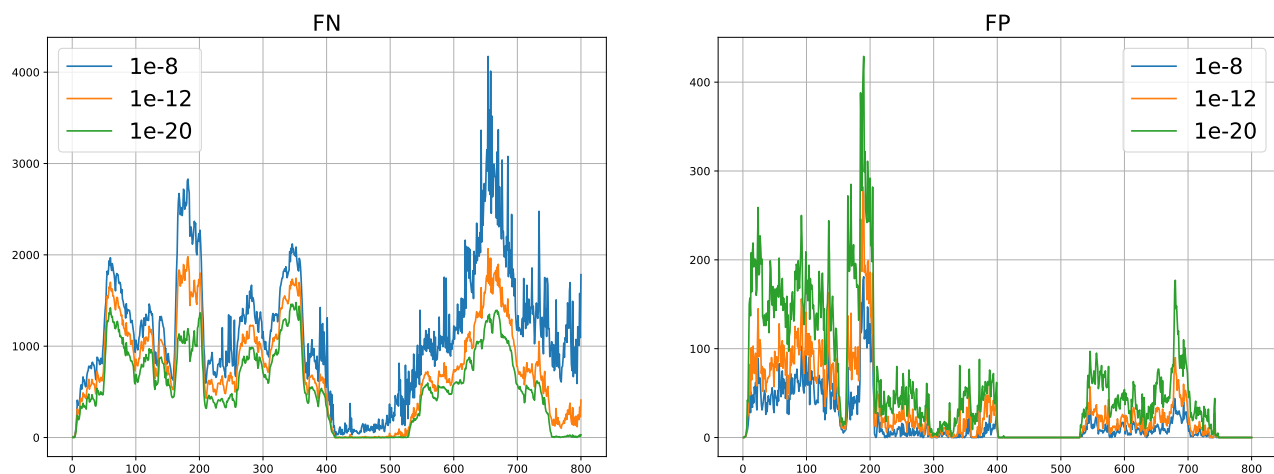


Рис. 9: Графики FN и FP для разных значений k

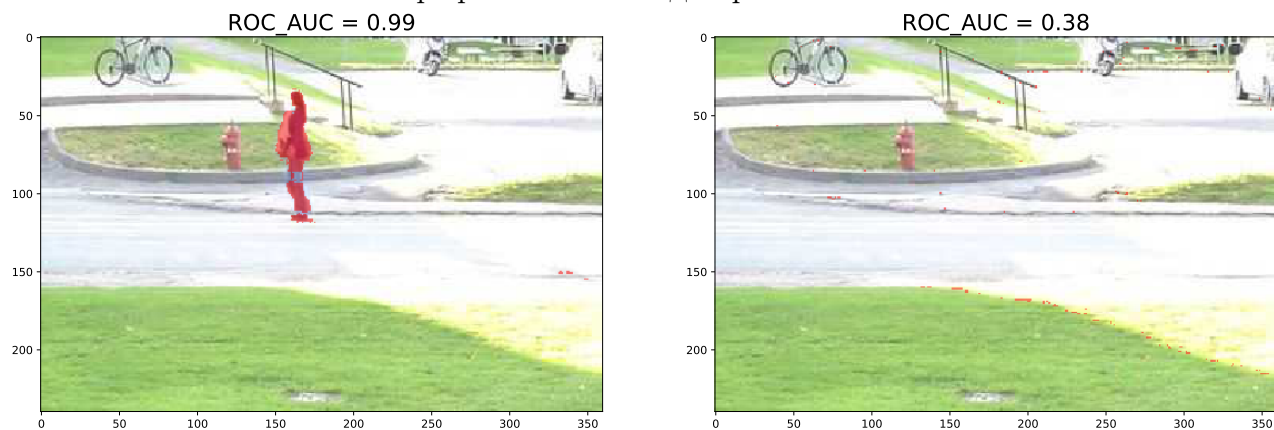


Рис. 10: Лучший и худший кадры в третьей модели
ROC-кривые для разных моментов времени

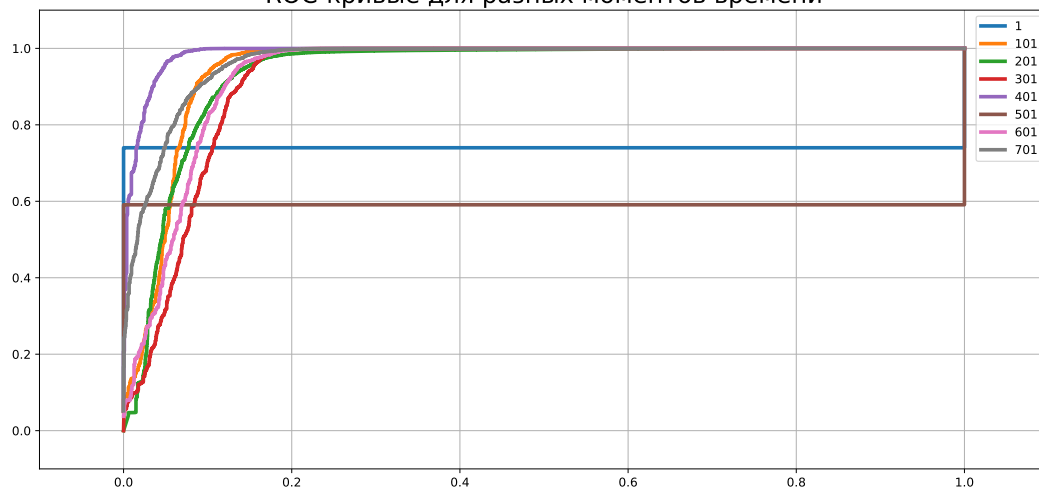


Рис. 11: Графики ROC-кривых для смеси гауссиан