ИНФОРМАЦИОННЫЕ И ТЕЛЕКОММУНИКАЦИОННЫЕ СЕТИ И СИСТЕМЫ

УДК 004.434

ПРОГРАММНО-АППАРАТНЫЙ МОДУЛЬ РАСПОЗНАВАНИЯ РУКОПИСНЫХ ОБРАЗОВ

Алимсеитова Жулдыз, *аспирант КГТУ им. И. Раззакова Кыргызской Республики +7 777 359 81 05, E-mail: zhuldyz_al@mail.ru*

В статье рассматривается программно-аппаратный модуль распознавания рукописных образов. Распознавание рукописных образов и применение их для аутентификации личности является сложной задачей, так как рукописный почерк является динамическим биометрическим параметром и меняется в связи с моральным и физическим состоянием человека. Кроме того, при распознавании рукописных образов используется несколько характеристик, например, степень нажима, наклон почерка, координаты начала, скорость написания. Исходя из этого в разработанном программно-аппаратном модуле используется эмулятор искусственной нейронной сети, позволяющий интеллектуализировать систему. В модуле реализованы режимы обучения и тестирования.

Ключевые слова: нейронная сеть, обучение системы, пароль, биометрический образ, тестирование системы, вероятность ошибки.

SOFTWARE AND HARDWARE MODULE OF THE HANDWRITING RECOGNITION IMAGES

Alimseitova Zhuldyz, *the graduate student of KGTU of I. Razzakov of the Kyrgyz Republic +7 777 359 81 05, E-mail: zhuldyz_al@mail.ru*

The article discusses the software and hardware module of the handwriting recognition images. Recognition of handwritten characters and their application for person authentication is a challenging task because handwriting handwriting is a dynamic biometric parameter is changing due to the moral and physical condition of the person. In addition, in recognition of handwritten images using multiple characteristics, for example, the degree of pressure, the slope of handwriting, the coordinates of the beginning, the speed of writing. Therefore, the developed hardware and software module is used, the emulator neural network, allowing to intellectualize the system. This module incorporates the modes of learning and testing.

Keywords: neural network training system, a password, a biometric image, system testing, the probability of error.

Необходимым условием выработки навыков стабильного написания рекомендуемого слова и корректного формирования обезличенной базы биометрических образов является освоение донором биометрии специализированного программно-аппаратного комплекса «Нейро-Тест 1.2» [3]. Он предназначен для проведения научных исследований, самостоятельного и под контролем инструктора выполнения процедур обучения и тестирования средства биометрико-нейросетевой аутентификации личности по рукописному слову-паролю, выполненного в соответствии с требованиями ГОСТ Р 52633.0–2006.

В программе используется эмулятор искусственной нейронной сети, имеющей множество выходов. Число выходов искусственной нейронной сети определяется длиной

порождаемого ею биометрического ключа. Это исключает взлом программы через обнаружение и подмену последнего бита решающего правила [2]. Программа имеет многобитовое решающее правило, сочетание значений бит которого уникально и злоумышленнику неизвестно. В программе «НейроТест 1.2» использован алгоритм быстрого автоматического обучения искусственной нейронной сети с 256 выходами.

Режим обучения. Обучение начинается с инициализации режима обучения (кнопка «Обучить» или «Режим «Обучение системы»). Так как для обучения система должна “знать” пароль (ключ) донора, необходимо сообщить системе обучающий пароль, то есть пароль, который после завершения обучения будет возникать при вводе правильного рукописного образа. Для этого необходимо выбрать в пункте главного меню "Режим" пункт "Задать пароль" (рисунок 1).

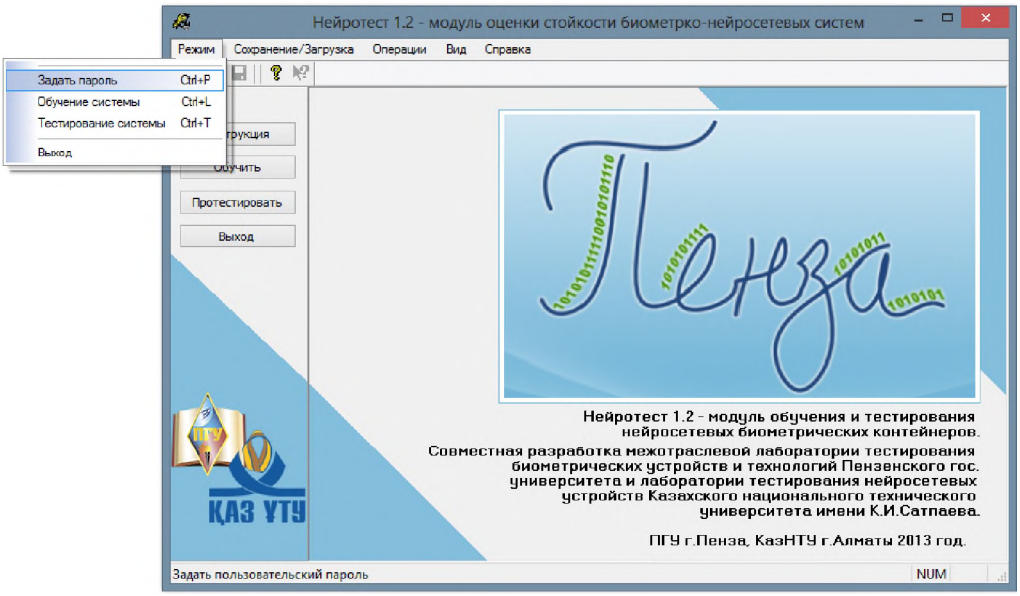

Рисунок 1 - Регистрация нового пользователя

В результате появиться окно где необходимо внести пользователя и пароль либо внести пользователя и нажать на кнопку "Автоматически сгенерировать новый пароль" (рисунок 2). Пароль может быть длиной от 1 до 32 символов. Если пользователь зарегистрирован в программе, то необходимо внести пользователя и при нажатии на кнопку загрузить пароль.

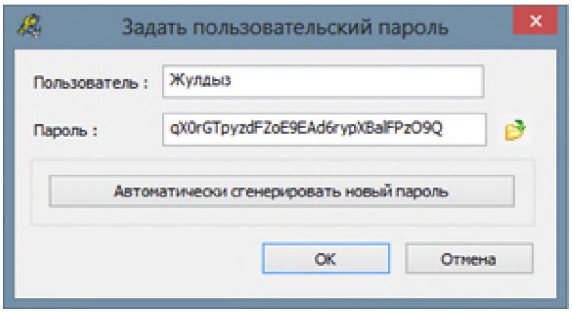

Рисунок 2 - Окно для регистрации пользователя и задания пароля

При нажатии на кнопку "ОК" появиться окно с запросом использовать или нет введенный пароль при обучении. При нажатии на кнопку "Да" выходит сообщение о том, что

можно приступить к обучению системы.

Далее необходимо воспроизвести своей рукой на поле графического планшета Genius G-Pen F350 (рисунок 3) заранее заданное рукописное слово.

Рисунок 3 – Графический планшет Genius G-Pen F350

После ввода рукописного слова нажать кнопку “Добавить”, при этом линованное поле очищается, а в правой части окна появляется номер очередного введенного примера. Если при вводе рукописного слова-пароля дрогнула рука, или образ записи не характерен, нажать кнопку “Очистить”. При этом слово удаляется без занесения в базу примеров. Необходимо воспроизвести слово 20 раз.

Можно просмотреть сохраненные примеры, щелкнув мышкой на интересующем номере в правой части окна (рисунок 4).

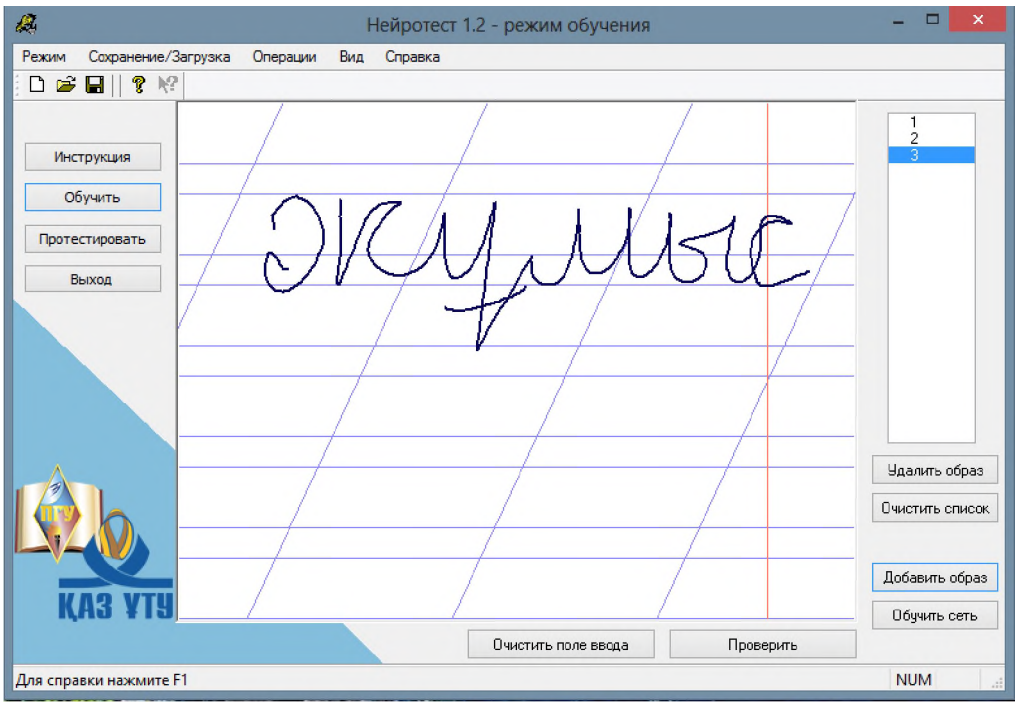

Рисунок 4 – Просмотр рукописных образов

Если какой-либо пример не соответствует почерку, то его необходимо удалить, нажав на кнопку «Удалить образ». Если не устраивают все образы или необходимо сменить слово-пароль, то удаляются все образы из списка (кнопка «Удалить всё»).

Обучение системы (нейросети) осуществляется нажатием кнопки «Обучить». При этом за время до 30 секунд появляется окно с прогнозом вероятностей ошибок системы и номером группы, к которой Вы относитесь (рисунок 5).

Если не устраивает номер группы и вероятностные характеристики, можно самостоятельно изменить номер группы и переучить нейросеть. Важно помнить, что не рекомендуется изменять номер группы более чем на 1 – 2 позиции, то есть можно переходить только в соседние позиции.

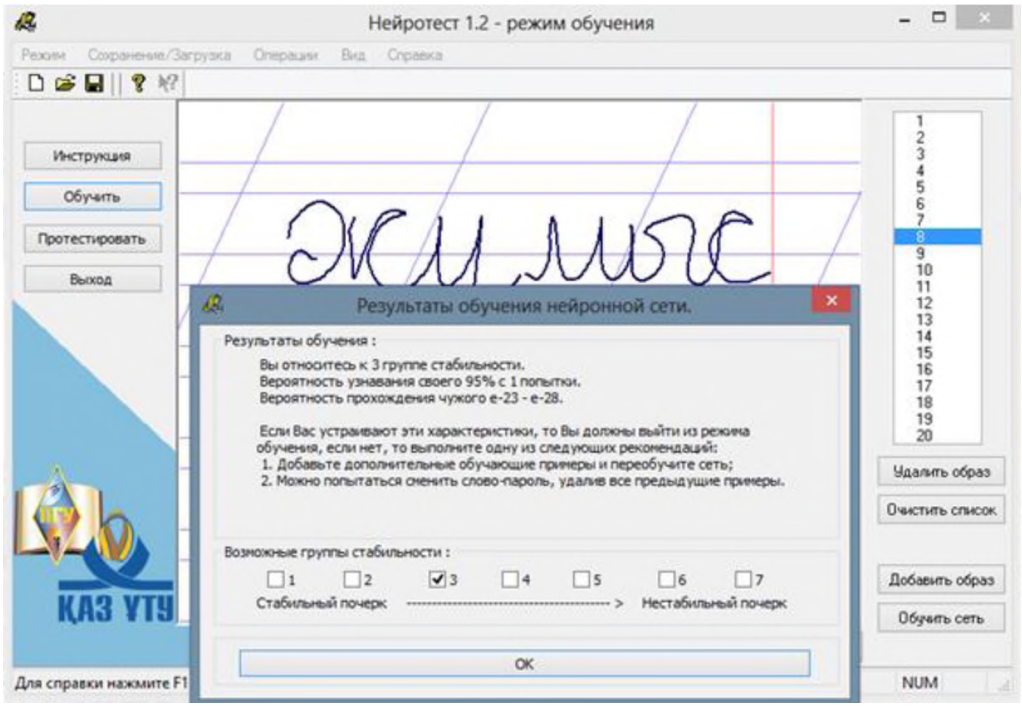

Рисунок 5 – Завершение обучения нейронной сети

Возможно, пользователь попал в седьмую группу, для которой вероятностные характеристики системы слишком плохи. Можно изменить условия обучения, удалив наиболее непохожий пример. Если имеются затруднения с выбором наихудшего примера, можно удалить все обучающие примеры и попробовать написать их заново, или добавить ещё несколько обучающих примеров. После удаления наихудшего или добавления дополнительного примера вновь нажать кнопку «Обучить». Если по-прежнему попасть в желаемую группу не удается, то удаляется самый непохожий пример или добавляется ещё один дополнительный и снова обучить сеть. Если после нескольких попыток повторно обучить сеть не удается попасть в желаемую группу стойкости, необходимо попытаться сменить слово.

Контроль распознавания «Своего». После обучения нейросети необходимо проверить качество узнавания системой «Своего». Проверка обучения осуществляется путем воспроизведения рукописного слова и нажатием кнопки «Проверить». При этом появляется окно (рисунок 6), показывающее сгенерированный нейросетью ключ в двоичной и шестнадцатеричной кодировках (звёздочками помечаются не совпавшие биты исходного ключа и сгенерированного). Также отображается общее количество не совпавших символов двоичного ключа.

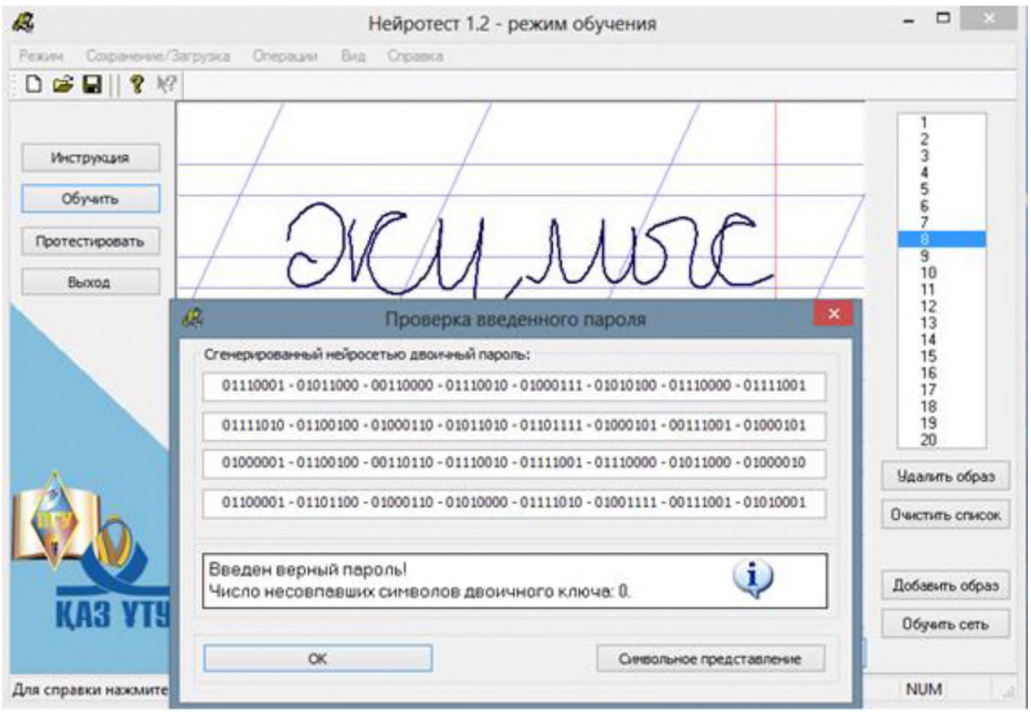

Рисунок 6 – Ввод образа «Свой»

Эти данные могут быть использованы для самостоятельного статистического тестирования системы.

Если система плохо узнает образ «Свой», то необходимо добавить нераспознанный образ в базу обучающих примеров и обучить сеть заново. После добавления нескольких новых образов сеть будет лучше узнавать вводящего рукописный образ, однако некоторые характеристики системы могут ухудшиться.

Если результаты обучения устраивают (ГРУППА 3…4), то необходимо сохранить данные на диск. Номер папки соответствует номеру персонального идентификатора донора биометрии.

Тестирование системы. Тестирование системы начинается с инициализации режима тестирования (кнопка «Протестировать» или «Режим >> Тестирование сети».

Прежде чем приступать к тестированию необходимо убедитесь, что нейросеть обучена. В режиме тестирования можно проверять вводимые рукописные образы на обученной сети, также можно добавить «удачный» образ в базу обучающих примеров.

Тестирование системы осуществляется путем воспроизведения рукописного слова-пароля и нажатием кнопки «Проверить введенный пароль». При этом загорается светофор в верхнем правом углу. Красный свет светофора соответствует очень большому расхождению динамики воспроизведения рукописного слова-пароля и вновь введенного проверочного слова (рисунок 7).

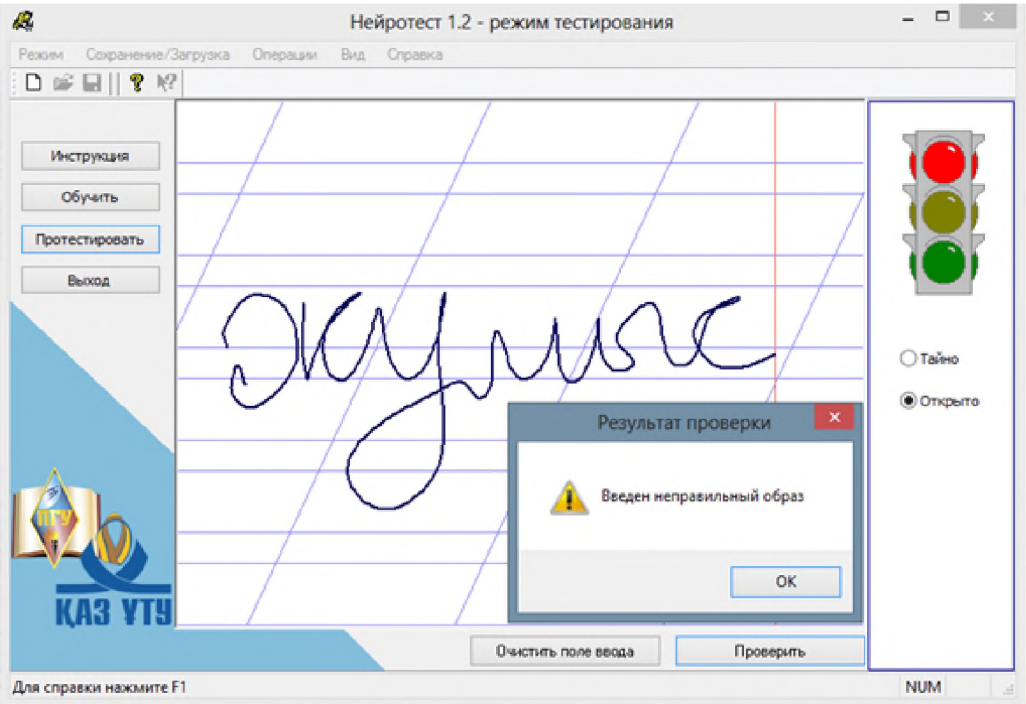

Рисунок 7 – Пример ввода неверного рукописного пароля

Частое загорание красного света в проверочном режиме при предъявлении «Своего» рукописного слова свидетельствует о плохой узнаваемости. Желтый свет загорается при несовпадении нескольких бит ключа, образ близок к эталонному. Зеленый свет соответствует полному совпадению введенного образа с эталонным (ключ воспроизводится нейросетью без ошибок) (рисунок 8).

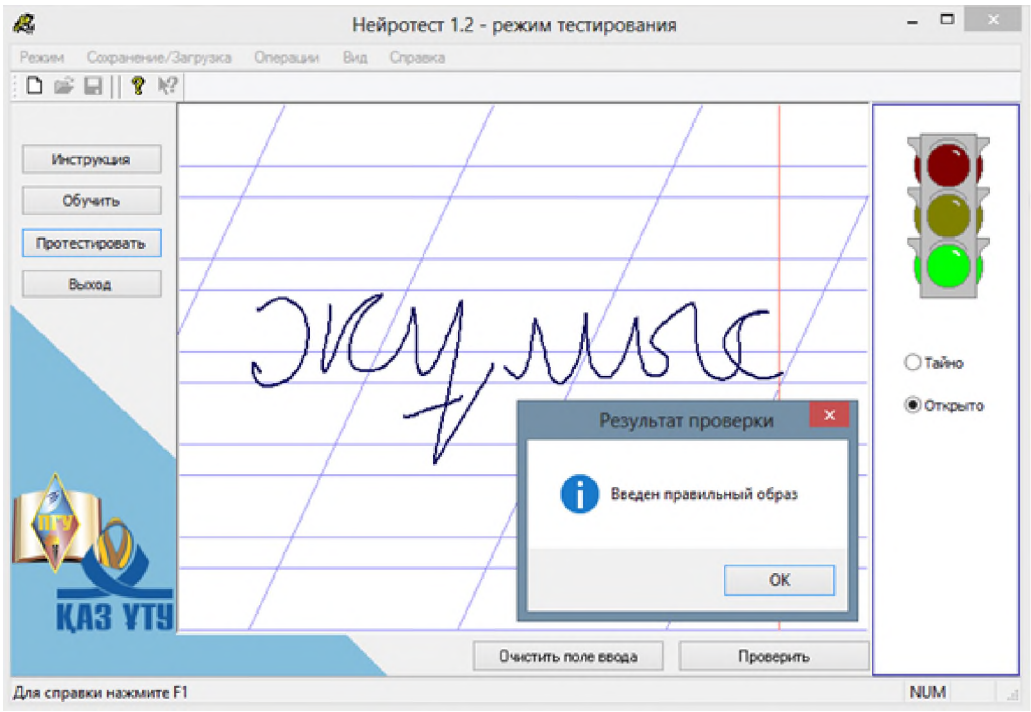

Рисунок 8 – Ввод правильного рукописного пароля

Самостоятельное статистическое тестирование системы. Данные по ошибкам первого и второго рода, приводимые в программе при обучении, являются результатом нейросетевого прогнозирования [2]. Очевидно, что любой прогноз нуждается в проверке. Возможно, самостоятельно проверить стойкость программы на своем почерке и на выбранном слове-пароле.

Оценка вероятности ошибок первого рода (отказ «Своему»). В режиме тестирования необходимо несколько раз написать выбранное ранее слово, фиксируя цвета светофора. Вероятность ошибки первого рода оценивается как отношение красных и желтых реакций светофора к общему числу попыток.

Оценка вероятности ошибок второго рода (пропуска «Чужого», не знающего пароль). В режиме тестирования необходимо воспроизводить случайные слова, нажимая клавишу “Проверить введенный пароль”. При этом фиксировать число не совпавших бит ключа. Среднее значение числа ошибок должно составить примерно 128. Далее нужно вычислить среднеквадратическое отклонение по формуле:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(128-x_i)^2}{N}} \qquad 1,$$

где N – общее число испытаний, x_i – число не совпавших бит в каждом испытании.

Стойкость может быть оценена по таблице 1. Эта таблица рассчитана для рукописного слова из 5 букв. При меньшей длине слова стойкость падает, при большем – увеличивается.

Таблица 1 - Взаимосвязь среднеквадратического отклонения со стойкостью системы

σ	5	10	15	20	25	30	35	40	45	50	55	60
P_2	10^{-120}	10^{-32}	10^{-15}	10^{-9}	10^{-6}	10^{-4}	10^{-3}	10^{-3}	10^{-2}	10^{-2}	10^{-2}	10^{-2}

Автоматизированное тестирование. В данной системе аутентификации предусмотрено два вида автоматизированного тестирования:

− на тестовых образах – тестирование на реальных рукописных образах из базы тестовых примеров. Данный вид тестирования позволяет оценить стойкость системы биометрической аутентификации при атаках с помощью больших баз биометрических примеров. Моделируется ситуация когда злоумышленник, располагающий базой образов, пытается взломать вашу защиту.

− на белом шуме – тестирование на искусственно синтезированных образах. Позволяет оценить стойкость системы к атакам машинного перебора. Моделируется ситуация, когда злоумышленник, выявив основные параметры биометрических коэффициентов, начинает подбирать параметры Вашего рукописного пароля.

Для запуска процедуры тестирования на тестовых образах следует выбрать пункт меню «**Операции**» “**Тестировать на тестовых образах**”. После того как пользователь укажет базу тестовых образов, запустится механизм оценки стойкости. Время тестирования существенно зависит от объёма базы биометрических образов. На рисунке 9 представлен полученный результат тестирования.

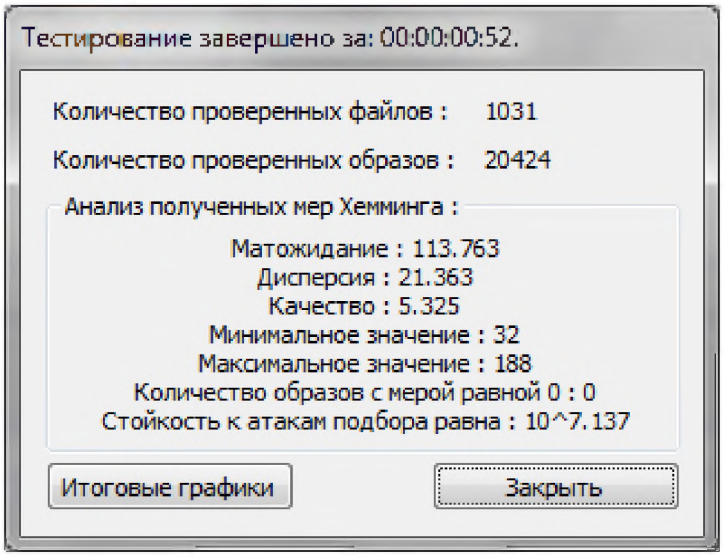

Рисунок 9 – Результат тестирования на тестовых образах

Для запуска процедуры тестирования на белом шуме следует выбрать пункт меню «**Операции**» “**Тестировать на белом шуме**”. Выдаваемый системой результат представлен на рисунке 10.

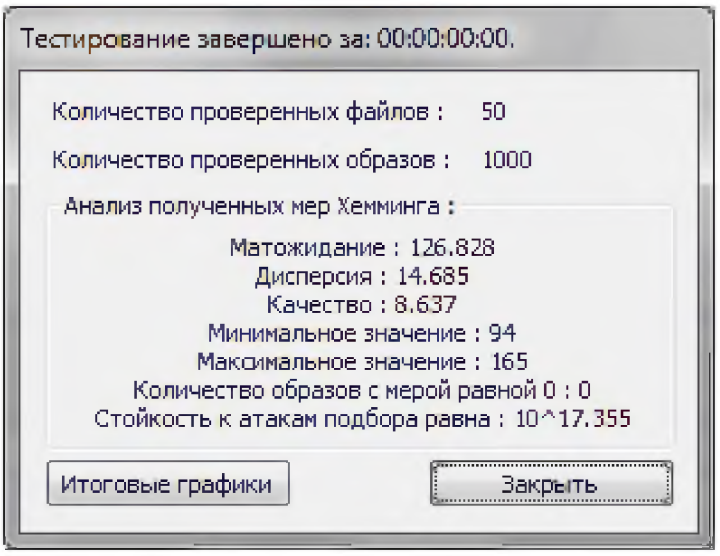

Рисунок 10 – Результат тестирования на белом шуме

Благодаря данным операциям пользователь может узнать распределение мер Хемминга (количество несовпадений бит ключа) и реальную стойкость только что обученной системы аутентификации. У идеально обученной системы математическое ожидание мер Хемминга равно 128, дисперсия равна 16, а качество – 8.

При низкой стойкости системы (ниже 10^{12}) рекомендуется переучить систему, удалив неудачные примеры, либо сменить используемое при обучении слово-пароль.

Вывод: Программно-аппаратный модуль распознавания рукописных образов позволяет обучить нейронную сеть на выбранном пользователем слове. Кроме того дается возможность удаления неудачных примеров и добавления новых для повышения группы стабильности. Также имеется несколько режимов тестирования, которые позволяют реально

оценить результаты обучения и стойкость системы.

Список литературы

1. ГОСТ Р 52633.0–2006 Защита информации. Техника защиты информации. Требования к средствам высоконадежной биометрической аутентификации.

2. Ахметов Б.С., Иванов А.И., Малыгин А.Ю., Фунтиков В.А. Основы биометрической аутентификации личности. Алматы: КазНТУ, 2014.

3. Ахметов Б.С., Алимсеитова Ж.К., Малыгин А.Ю., Юбузова Х.И. Формирование биометрической базы рукописных образов на казахском языке для программ биометрической аутентификации личности. Труды II международной научно-практической конференции «Информационные и телекоммуникационные технологии: образование, наука, практика» - Алматы: КазНТУ, 2015. – Том 2 -с. 32-35