

Взаимосвязь нейросетевых оценок эмоциональных состояний пилота с качеством решения функциональных задач

О.Н. Корсун, В.Н. Юрко

Аннотация — В работе представлены результаты анализа диагностических данных пилотов, а именно зарегистрированных эмоциональных состояний во время проведения следующих экспериментов: выполнения многократных заходов на посадку по директорным стрелкам на пилотажно-навигационном приборе; выполнения задачи поиска объектов. Для определения эмоциональных состояний применялись глубокие сверточные нейронные сети. При анализе проводимых экспериментов использовались методы математической статистики, метод главных компонент для уменьшения размерности матрицы данных по 7 исследуемым эмоциональным категориям с наименьшими потерями информации. В экспериментах по заходу на посадку оценивалась зависимость среднего количества эмоций определенной категории в минуту от среднеквадратического отклонения от линии глиссады по вертикали. Проводился анализ данных по всему заходу. Также оценивалась зависимость среднего количества эмоций данной категории в начале захода от среднего суммарного отклонения от линии глиссады при приближении к взлетно-посадочной полосе в целях прогнозирования точности захода на посадку. В экспериментах по поиску пилотами объектов оценивались вероятности регистрации эмоций от времени поиска объектов. В процессе анализа данных было определено, что эмоциональная категория «Печаль (Sad)» является признаком сосредоточенности на выполнении задачи, признаком концентрации внимания пилота. Эмоциональные категории «Страх (Fear)» и «Удивление (Surprise)» являются признаками растерянности и обеспокоенности. Построены обобщенные корреляционные матрицы, которые показали, что эмоции коррелированы друг с другом, с показателями точности пилотирования (отклонениями от линии глиссады), временем реализации (временем поиска объектов).

Ключевые слова — полунатурное моделирование, эмоциональные состояния, нейронные сети, логнормальное распределение.

I. ВВЕДЕНИЕ

Увеличение сложности современных авиационных полунатурных комплексов приводит к возрастанию роли тренажерной подготовки летных экипажей. Отработка взаимодействия с информационно-управляющим полем кабины, между членами экипажа на различных этапах полета, отработка действий в особых режимах полета, воспроизведение которых в реальных условиях

невозможно или связано с высоким риском, - важнейшие элементы наземной подготовки не только при обучении, но и при восстановлении и поддержании навыков опытных летчиков[1-3].

II. ОЦЕНКА ЭМОЦИОНАЛЬНЫХ СОСТОЯНИЙ ПИЛОТА ВО ВРЕМЯ ВЫПОЛНЕНИЯ МНОГОКРАТНЫХ ЗАХОДОВ НА ПОСАДКУ

В настоящее время глубокие сверточные нейронные сети считаются наиболее перспективными для задач распознавания образов[2-4], в том числе эмоциональных состояний операторов (рис. 1).

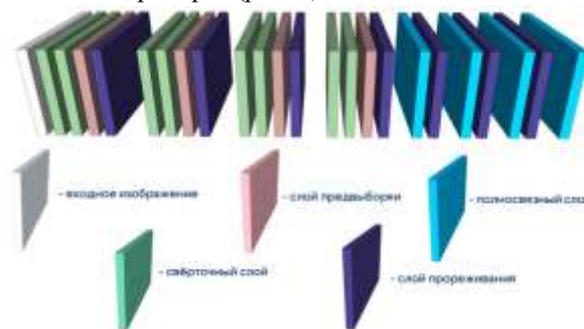


Рис. 1. Структура глубокой нейронной сети (3D-модель слоев)

В данном исследовании при разработке специального программного обеспечения для обучения нейронной сети был выбран набор данных FER-2013 (<https://www.kaggle.com/datasets/msambare/fer2013>), который состоит из 35887 монохромных изображений размером 48×48 пикселей с 7 типами эмоций (гнев, отвращение, страх, радость, грусть, удивление, нейтральное состояние). Стоит отметить, что в мире не существует эталонного выражения той или иной эмоции. При формировании баз данных с разметкой лиц по эмоциональным категориям среди специалистов есть несогласованность в том, какую эмоцию они видят. Для детектирования контуров лиц на изображении выбран каскад Хаара – один из классических способов распознавания классов объектов с большой скоростью работы.

Классический метод обучения нейронной сети – метод минимизации эмпирического риска – нахождение в заданной модели A алгоритма a , доставляющего

минимальное значение функционалу качества Q на заданной обучающей выборке.

Функционал качества (эмпирический риск) алгоритма a на заданной обучающей выборке X_{learn} :

$$Q(a, X_{learn}) = \frac{1}{k} \sum_{i=1}^k L(a, x_i),$$

где $L(a, x)$ - функция потерь, характеризующая величину ошибки алгоритма a на объекте x.

Проверка предложенной структуры нейронной сети на тестовом наборе, содержащем 3589 изображений, показала 79,35 % правильных распознаваний [4-11].

Результаты анализа видеоизображения лица пилота на предмет распознавания эмоциональных состояний пилота во время выполнения многократных заходов на посадку представлены ниже. Посадка – это наиболее сложный этап полета.

На рис. 2 представлена зависимость среднего количества эмоций «Печаль (Sad)» в минуту от СКО отклонения от линии глиссады по вертикали. Проводился анализ данных по всему заходу. Красная точка на графике – один из множества заходов на посадку.

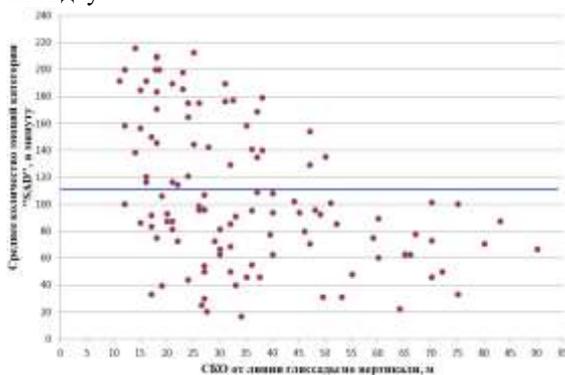


Рис. 2. График зависимости среднего количества эмоций «Печаль (Sad)» от СКО отклонения от линии глиссады по вертикали

Для проверки согласия (goodness-of-fit) эмпирического распределения значений СКО отклонения от линии глиссады с теоретическим (логнормальным) выполнялся тест Хи-квадрат (χ^2 -тест). Тест сравнивает наблюдаемые частоты с ожидаемыми, оценивая p-value: если $p > 0.05$, гипотеза о соответствиях распределений отклоняется. В нашем случае (рис.3) $p = 0.52$ позволяет сделать вывод о том, что распределение данных существенно не отличается от теоретического (логнормального).

Случайная величина X имеет логнормальное распределение с параметрами (m, σ), если ее плотность вероятности определяется выражением [12]:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - m)^2}{2\sigma^2}\right)$$

Вид плотности вероятности логнормального распределения при $m=3.42, \sigma=0.5$ представлен на рис. 3.

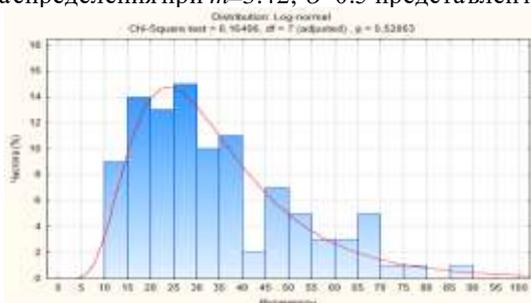


Рис. 3. Подбор теоретического логнормального распределения для значений СКО во всем диапазоне значений регистрируемых эмоций «Печаль(Sad)»

В данном эксперименте (анализируя движение при приближении к глиссаде и по глиссаде) высокая средняя точность по отклонению от глиссады соответствует 40 метрам. Стоит задача определения значений диагностического признака - значений регистрируемых эмоций «Печаль (Sad)» при котором вероятность выполнения захода на посадку с высокой точностью выше. Для этого разграничим данные, представленные на рисунке 2, на две части (условная синяя линия) и выполним подгонку логнормального распределения для значений СКО.

На рис. 4 – подбор теоретического логнормального распределения для значений СКО в полосе значений диагностического признака от 0 до 110.

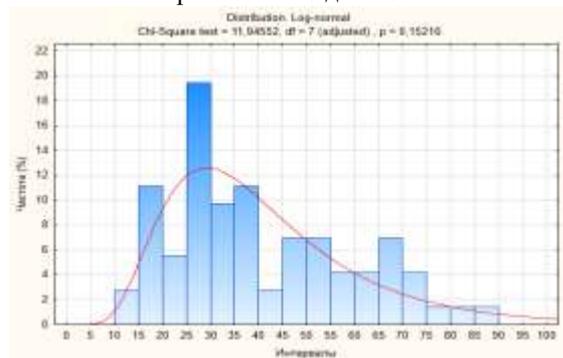


Рис. 4. Подбор логнормального распределения для значений СКО в полосе значений диагностического признака «Печаль (Sad)» от 0 до 110

На рис. 5 - подбор теоретического логнормального распределения для значений СКО в полосе значений диагностического признака от 110 до 220.

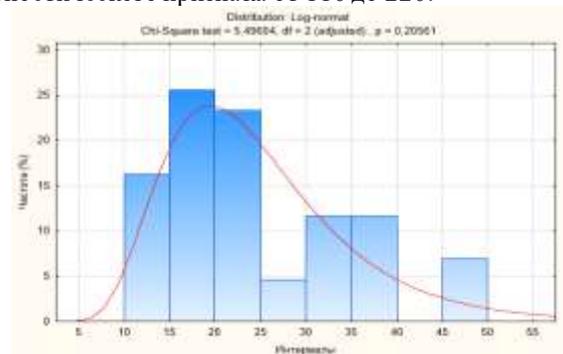


Рис. 5. Подбор логнормального распределения для значений СКО в полосе значений диагностического признака «Печаль (Sad)» от 110 до 220

Ориентируясь на среднюю точность по отклонению от глиссады в 40 метров при количестве эмоций «Печаль (Sad)» от 0 до 110 в 56% случаев пилот выполняет задачу с высокой точностью. При количестве эмоций от 110 до 220 - в 93% случаев пилот выполняет задачу с высокой точностью.

Интерес также представляет прогнозирование точности захода на посадку. На рис. 6 представлен график зависимости среднего количества эмоций «Печаль (Sad)» в начале захода от среднего суммарного отклонения от линии глиссады при приближении к взлетно-посадочной полосе (ВПП).

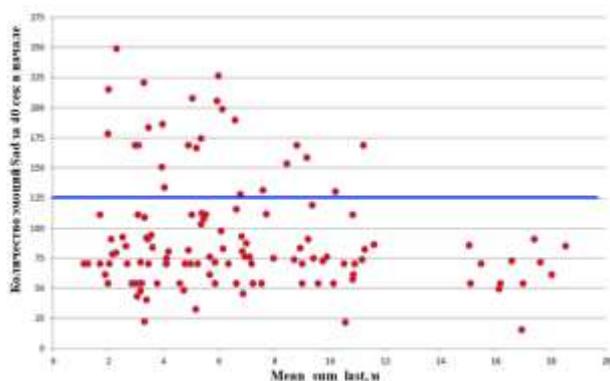


Рис. 6. График зависимости среднего количества эмоций «Печаль (Sad)» в начале захода от среднего суммарного отклонения от линии глиссады при приближении к ВПП

Ориентируясь на точность по отклонению от глиссады в 10 метров при количестве эмоций «Печаль (Sad)» от 0 до 125 в 77% случаев пилот выполняет задачу с высокой точностью. При количестве эмоций от 125 до 250 - в 94% случаев пилот выполняет задачу с высокой точностью.

III. ОЦЕНКА ЭМОЦИОНАЛЬНЫХ СОСТОЯНИЙ ПИЛОТА ВО ВРЕМЯ ВЫПОЛНЕНИЯ ЗАДАЧИ ПОИСКА ОБЪЕКТА

Задача поиска возникает тогда, когда требуется определить положение некоторого объекта, находящегося в заданной области пространства с помощью поисковых средств. Поисковая операция должна обеспечить обнаружение объекта за минимальное время. В данной работе приводятся результаты экспериментов на полунатурном комплексе по обнаружению пилотом объекта, находящегося на земной поверхности. В каждой серии выполнялось до 26 однотипных реализаций по поиску объектов. В начале эксперимента пилот в течение 3-5 секунд смотрит на пилотажно-навигационный дисплей для фиксации исходного углового положения головы, а затем поворачивает голову и выполняет визуальный поиск объекта, находящегося на земле. Для имитации внекабинной обстановки на полунатурном комплексе разработана система визуализации внекабинного пространства [13].

Время реализации (поиска объекта) определялось как интервал между нажатием кнопки 1 в макете кабины полунатурного стенда в момент начала поиска, и нажатием кнопки 2 при совмещении монокла на шлемной системы, связанного с положением головы, с объектом поиска (окончание поиска).

В процессе эксперимента все параметры полета поступали в систему бортовой регистрации, кроме того, производилась видеозапись лица пилота. При обработке результатов анализ видеозображения лица для распознавания эмоциональных состояний проводился в течение 2с до момента нажатия кнопки 2 [14-18].

Уменьшение времени реализации характеризуется ростом числа эмоциональных состояний типа «Sad (Печаль)» (рис. 7), что соответствует повышению степени концентрации внимания [14].

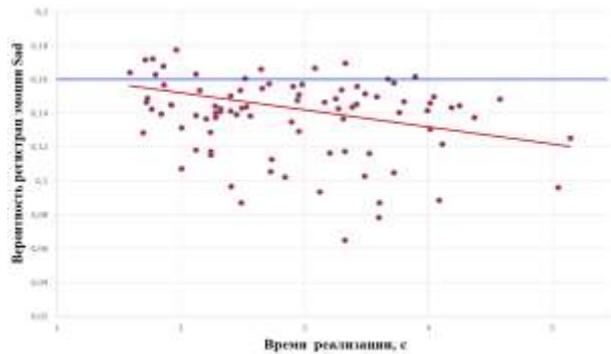


Рис. 7. График зависимости вероятности регистрации эмоции категории «Печаль (Sad)» от времени реализации для 4 серий экспериментов

На рис. 8 – подбор теоретического логнормального распределения для значений времени реализации в полосе значений диагностического признака «Печаль (Sad)» от 0 до 0,16.

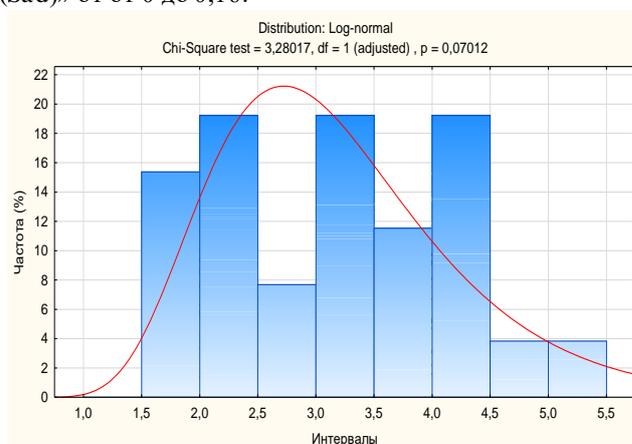


Рис. 8. Подбор логнормального распределения для значений времени реализации в полосе значений диагностического признака от 0 до 0,16

На рис. 9 – подбор теоретического логнормального распределения для значений времени реализации в полосе значений диагностического признака «Печаль (Sad)» от 0,16 до 0,2.

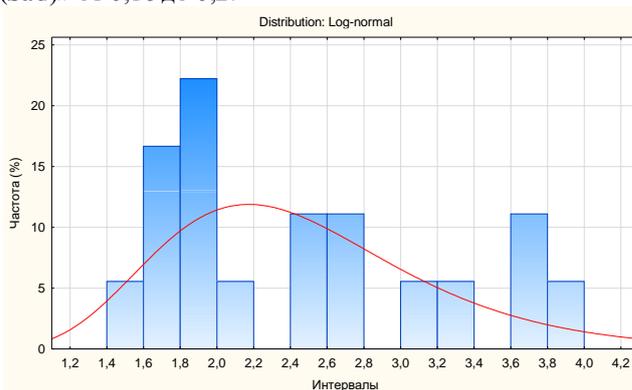


Рис. 9. Подбор логнормального распределения для значений времени реализации в полосе значений диагностического признака от 0,16 до 0,2

Ориентируясь на значение времени реализации (поиска) в 3,5 секунды при значениях вероятностей регистрации эмоции «Печаль (Sad)» от 0 до 0,16 в 67% случаев пилот выполняет задачу с высокой точностью. При количестве эмоций от 0,16 до 0,2 - в 91% случаев пилот выполняет задачу с высокой точностью.

На рис. 10 и 13 представлены зависимости по вероятностям регистрации эмоций категорий «Страх (Fear)» и «Удивление (Surprise)» от времени реализации. Данные эмоциональные категории являются признаками растерянности и обеспокоенности.

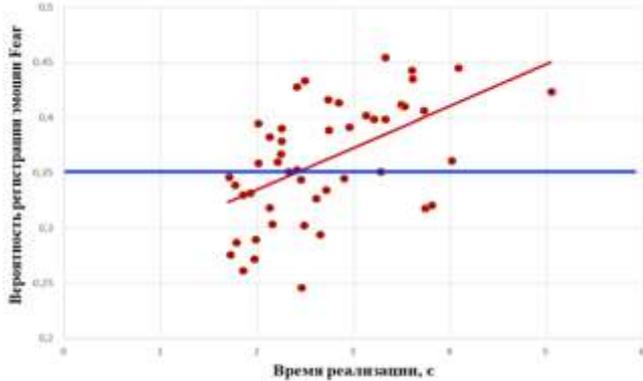


Рис. 10. График зависимости вероятности регистрации эмоции категории «Страх (Fear)» от времени реализации

На рис. 11 – подбор теоретического логнормального распределения для значений времени реализации в полосе значений диагностического признака «Страх (Fear)» от 0 до 0,35.

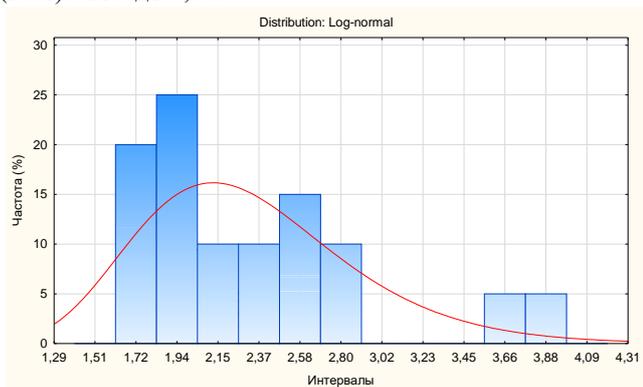


Рис. 11. Подбор логнормального распределения для значений времени реализации в полосе значений диагностического признака от 0 до 0,35

На рис. 12 – подбор теоретического логнормального распределения для значений времени реализации в полосе значений диагностического признака «Страх (Fear)» от 0,35 до 0,5.

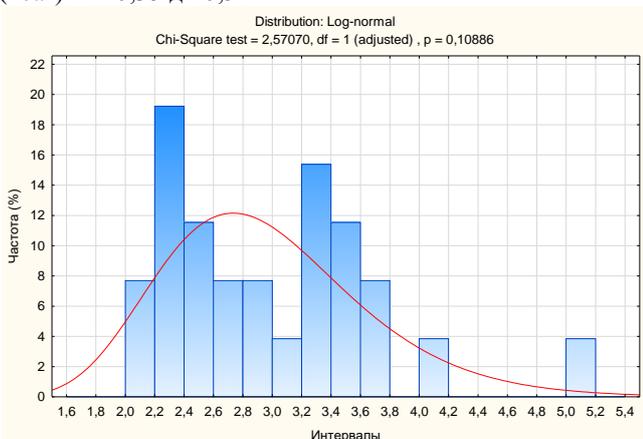


Рис. 12. Подбор логнормального распределения для значений времени реализации в полосе значений диагностического признака от 0,35 до 0,5

Ориентируясь на значение времени реализации (поиска) в 3,5 секунды при значениях вероятностей

регистрации эмоции «Страх (Fear)» от 0 до 0,35 в 97% случаев пилот выполняет задачу с высокой точностью. При количестве эмоций от 0,35 до 0,5 - в 78% случаев пилот выполняет задачу с высокой точностью.

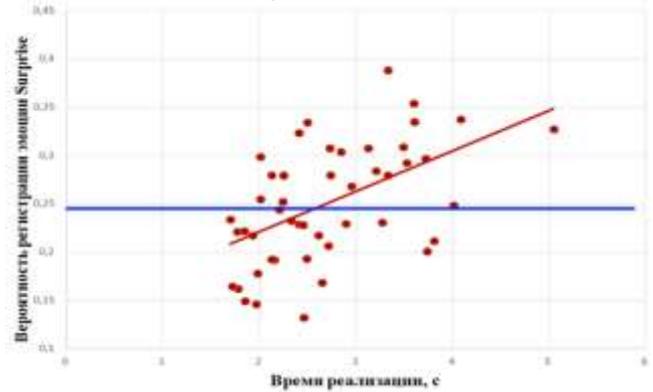


Рис. 13. График зависимости вероятности регистрации эмоции категории «Удивление(Surprise)» от времени реализации

На рис. 14 – подбор теоретического логнормального распределения для значений времени реализации в полосе значений диагностического признака «Удивление(Surprise)» от 0 до 0,24.

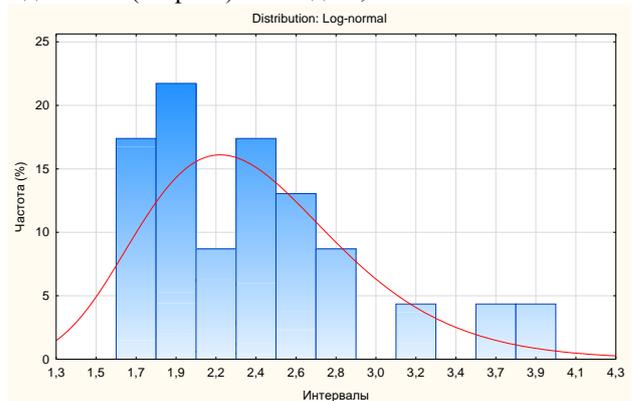


Рис. 14. Подбор логнормального распределения для значений времени реализации в полосе значений диагностического признака от 0 до 0,24

На рис. 15 – подбор теоретического логнормального распределения для значений времени реализации в полосе значений диагностического признака «Удивление(Surprise)» от 0,24 до 0,4.

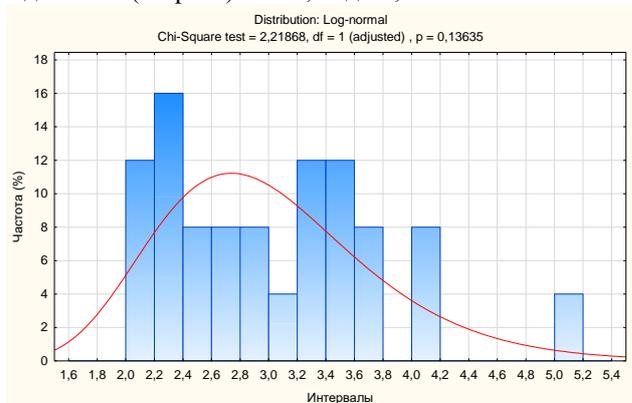


Рис. 15. Подбор логнормального распределения для значений времени реализации в полосе значений диагностического признака от 0,24 до 0,4

Ориентируясь на значение времени реализации (поиска) в 3,5 секунды при значениях вероятностей регистрации эмоции «Удивление(Surprise)» от 0 до 0,24 в 96% случаев пилот выполняет задачу с высокой

точностью. При количестве эмоций от 0,24 до 0,4 - в 76% случаев пилот выполняет задачу с высокой точностью.

Метод главных компонент применяется для снижения размерности пространства наблюдаемых векторов, не приводя к существенной потере информативности. Главные компоненты представляют собой ортогональную систему координат, в которой дисперсии характеризуют их статистические свойства.

Соотношения для определения всех p главных компонент матрицы X могут быть представлены в виде $Z = LX$, где $Z = (z^{(1)}, \dots, z^{(p)})'$, $X = (x^{(1)}, \dots, x^{(p)})'$, а матрица L состоит из строк $l_j = (l_{j1}, \dots, l_{jp})$, $j = \overline{1, p}$, являющихся собственными векторами ковариационной матрицы Σ случайной величины X . Этим собственным векторам соответствуют собственные числа λ_j , $j = \overline{1, p}$.

На рис. 16 представлены тренды по первым двум главным компонентам матрицы данных по 7 эмоциональным категориям от времени реализации в проводимых экспериментах. Информативна первая главная компонента, имеющая восходящий тренд.



Рис. 16. График зависимости первых двух главных компонент матрицы данных по 7 эмоциональным категориям от времени реализации

IV. ЗАКЛЮЧЕНИЕ

Проведен анализ видеоизображения лица пилотов на предмет распознавания эмоциональных состояний во время выполнения многократных заходов на посадку и проведения экспериментов по поиску объектов. Определены эмоциональные категории, которые связаны с точностью выполнения функциональных задач. Информативным показателем является эмоциональная категория «Печаль (Sad)», отображающая на лице оператора признаки сложного когнитивного процесса – концентрации, сосредоточения внимания. Эмоциональные категории «Удивление (Surprise)» и «Страх (Fear)» являются признаками растерянности и обеспокоенности. Определены конкретные значения диагностических признаков пилотов в целях прогнозирования выполнения поставленных задач с заданной точностью.

БИБЛИОГРАФИЯ

[1] Korsun O. N., Yurko V. N. and Mikhaylov E. I. Operator's State Estimation Based on the Face's Video Images Analysis Using Deep Convolutional Neural Networks. IOP Conf. Series: Materials Science and Engineering 714 (2020) 012012 doi:10.1088/1757-899X/714/1/012012.

[2] Желтов С.Ю., Себряков Г.Г. Моделирование многофункциональной деятельности человека-оператора при управлении летательным аппаратом // XXVII Всероссийская научно-техническая конференция школы-семинара «Передача,

прием, обработка и отображение информации о быстропротекающих процессах»: сборник статей. 2016. С. 266–275.

[3] Современные аудиотехнологии в информационно-управляющем поле кабины пилотов / Г. Г. Себряков, О. Н. Корсун, Г. А. Лаврова, А. О. Лавров, А. В. Полиев, В. Н. Юрко — Москва : Издательский дом Академии имени Н. Е. Жуковского, 2021.

[4] Korsun O.N., Yurko V.N., Mikhaylov E.I. Operator's State Estimation Based on the Face's Video Images Analysis Using Deep Convolutional Neural Networks. IOP Conf. Series: Materials Science and Engineering. IOP Publishing, 2020. doi: 10.1088/1757-899X/714/1/012012

[5] Хисматов И.Ф., Рожков К.Ю. Вопросы радиоэлектроники. Метод коррекции изображений неподвижных объектов в авиационных системах наблюдения // Серия: Техника телевидения. 2014. № 2. С. 126-133.

[6] Моисеев Н.Н. Численные методы в теории оптимальных систем. М.: Наука, 1971. Вандер Плас Дж. Python для сложных задач: наука о данных и машинное обучение. СПб.: Питер. 2022. С. 200–223.

[7] Овчинникова Т.В., Корсун О.Н., Кувшинов В.С. Методика оценки адекватности результатов математического моделирования и натуральных экспериментов // В сборнике: Передача, приём, обработка и отображение информации о быстропротекающих процессах. материалы XXV Юбилейной Всероссийской научно-технической конференции школы-семинара. Российская академия ракетных и артиллерийских наук 2014. С. 513-520.

[8] Авиация ВВС России и научно-технический прогресс / Под ред. Е.А.Федосова, М.: «Дрофа», 2005

[9] Shapiro, L.G. and Stockman, G.C. (2001) Computer Vision. Prentice Hall, Upper Saddle River.

[10] Соловьев Ю.А. Спутниковые дифференциальные подсистемы посадки воздушных судов. // Новости навигации. 2014. №2. С. 20-29.

[11] Монсик В. Б., Скрынников А. А. Вероятность и статистика : учебное пособие. М. : Бином. Лаборатория знаний, 2011.

[12] Система визуализации закабинной обстановки и самолёта. Каширкин С.В., Реховский А.В., Звягин С.С., Сапронова А.А., Пичикин С.В., Морозов И.А., Пискарев К.Ю. Свидетельство о регистрации программы для ЭВМ RU 2018610856, 18.01.2018. Заявка № 2017662100 от 24.11.2017.

[13] Корсун О.Н., Юрко В.Н. Методы контроля психофизиологического состояния оператора по эмоциональному выражению лица и анализу характеристик морганий на основе глубоких сверточных нейронных сетей // Вестник Московского государственного технического университета им. Н.Э. Баумана. Серия Приборостроение. 2021. № 1 (134). С. 120-134.

[14] Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In ICLR, 2015.

[15] Ekman P., Wallace V.F. Unmasking the Face: A Guide to Recognizing Emotions From Facial Expressions. Malor Books, 2003.

[16] Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks // Advances in Neural Information Processing Systems. 2012. P. 1097–1105.

[17] Хайкин С. Нейронные сети: полный курс, 2-е издание. – Издательский дом Вильямс, 2008.

The relationship between neural network assessments of a pilot's emotional states and the quality of functional task performance

O.N. Korsun, V.N. Yurko

Abstract — The article presents the results of an analysis of diagnostic instrumental data from pilots, specifically focusing on recorded emotional states during the following experiments: performing multiple landing approaches using flight director bars on a flight navigation display; performing object search tasks. Deep convolutional neural networks were employed to identify emotional states. Mathematical statistics methods were used to analyze the experimental data, including principal component analysis to reduce the dimension of the data matrix across the 7 studied emotional categories while minimizing information loss. In the landing approach experiments, the relationship between the average number of emotions per minute for a specific category and the RMS (root mean square) deviation from the glideslope was evaluated. Data analysis was conducted throughout the entire approach. Furthermore, the relationship between the average number of emotions of a given category at the beginning of the approach and the average total deviation from the glideslope when nearing the runway was assessed, aiming to predict landing approach accuracy. In the object search experiments, the probabilities of registering emotions over the object search time were evaluated. Data analysis revealed that the emotional category "Sad" serves as an indicator of focus on task performance and a indicator of the pilot's concentration. The emotional categories "Fear" and "Surprise" were identified as indicators of confusion and concern. Generalized correlation matrices were also calculated, demonstrating that emotions are correlated with each other, with piloting precision metrics (deviations from the glideslope), and with task completion time (object search time).

Keywords — semi-natural (hardware-in-the-loop) simulation, emotional states, neural networks, lognormal distribution.

REFERENCES

- [1] Korsun O. N., Yurko V. N. and Mikhaylov E. I. Operator's State Estimation Based on the Face's Video Images Analysis Using Deep Convolutional Neural Networks. IOP Conf. Series: Materials Science and Engineering 714 (2020) 012012 doi:10.1088/1757-899X/714/1/012012.
- [2] Zheltov S.YU., Sebyakov G.G. Modelirovanie mnogofunkcionalnoj deyatel'nosti cheloveka-operatora pri upravlenii letatel'nym apparatom. In XXVII Vserossiyskaya nauchno-tekhnicheskaya konferenciya shkoly-seminara «Peredacha, priem, obrabotka i otobrazhenie informacii o bystroprotekayushchih processah»: sbornik statej. 2016. S. 266–275.
- [3] Sovremennye audiotekhnologii v informacionno-upravlyayushchem pole kabiny pilotov / G. G. Sebyakov, O. N. Korsun, G. A. Lavrova, A. O. Lavrov, A. V. Poliev, V. N. YURKO Moskva : Izdatel'skij dom Akademii imeni N. E. Zhukovskogo, 2021.
- [4] Korsun O.N., Yurko V.N., Mikhaylov E.I. Operator's State Estimation Based on the Face's Video Images Analysis Using Deep Convolutional Neural Networks. IOP Conf. Series: Materials Science and Engineering. IOP Publishing, 2020. doi: 10.1088/1757-899X/714/1/012012
- [5] Hismatov I.F., Rozhkov K.YU. Voprosy radioelektroniki. Metod korrekcii izobrazhenij nepodviznykh ob'ektov v aviacionnyh sistemah nablyudeniya. Seriya: Tekhnika televizeniya. 2014. № 2. S. 126-133.
- [6] Moiseev N.N. CHislennyye metody v teorii optimal'nyh sistem. M.: Nauka, 1971.
- [7] Vander Plas Dzh. Python dlya slozhnyh zadach: nauka o dannyh i mashinnoe obuchenie. SPb.: Piter. 2022. S. 200–223.
- [8] Ovchinnikova T.V., Korsun O.N., Kuvshinov V.S. Metodika ocenki adekvatnosti rezul'tatov matematicheskogo modelirovaniya i natumykh eksperimentov. In Peredacha, priyom, obrabotka i otobrazhenie informacii o bystroprotekayushchih processah. materialy HXV YUbilejnoy Vserossiyskoj nauchno-tekhnicheskoy konferencii shkoly-seminara. Rossiyskaya akademiya raketnyh i artillerijskih nauk. 2014. S. 513-520.
- [9] Aviatsiya VVS Rossii i nauchno-tekhnicheskij progress / Pod red. E.A.Fedosova, M.: «Drofa», 2005
- [10] Shapiro, L.G. and Stockman, G.C. (2001) Computer Vision. Prentice Hall, Upper Saddle River.
- [11] Solov'ev YU.A. Sputnikovye differencial'nye podsistemy posadki vozдушnyh sudov. Novosti navigacii. 2014. №2. S. 20-29.
- [12] Monsik V. B., Skrynnikov A. A. Veroyatnost' i statistika : uchebnoe posobie. M. : Binom. Laboratoriya znanij, 2011.
- [13] Sistema vizualizacii zakabinnoj obstanovki i samolyota. Kashirkin S.V., Rekhovskij A.V., Zvyagin S.S., Sapronova A.A., Pichikin S.V., Morozov I.A., Piskaryov K.YU. Svidetel'stvo o registracii programmy dlya EVM RU 2018610856, 18.01.2018. Zayavka № 2017662100 ot 24.11.2017.
- [14] Korsun O.N., Yurko V.N. Metody kontrolya psihofiziologicheskogo sostoyaniya operatora po emocional'nomu vyrazheniyu lica i analizu harakteristik morganiy na osnove glubokih svertochnykh nejronnyh setej. Vestnik Moskovskogo gosudarstvennogo tekhnicheskogo universiteta im. N.E. Baumana. Seriya Priborostroenie. 2021. № 1 (134). S. 120-134.
- [15] Simonyan K., Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In ICLR, 2015.
- [16] Ekman P., Wallace V.F. Unmasking the Face: A Guide to Recognizing Emotions From Facial Expressions. Malor Books, 2003..
- [17] Krizhevsky A., Sutskever I., Hinton G.E. Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems. 2012. P. 1097–1105.
- [18] Hajkin S. Nejronnye seti: polnyj kurs, 2-e izdanie. – Izdatel'skij dom Vil'yams, 2008. Hasan K.S., Ng V. Automatic Keyphrase Extraction: A Survey of the State of the Art. In 52nd Annual Meeting of the Association for Computational Linguistics, vol 1: Long Papers, pp. 1262–1273, 2014.