Artificial Intelligence in Healthcare: A New Frontier in Predicting and Managing Chronic Diseases

Natalie Nathan MD[1], Michael Saring MD[1], Noam Savion-Gaiger MD[1], Kira Radinsky PhD[1,2], and Alma Peri MD[1]

[1]Diagnostic Robotics, Tel Aviv, Israel
[2]Taub Faculty of Computer Science, Technion–Israel Institute of Technology, Haifa, Israel

ABSTRACT A rise in the incidence of chronic health conditions, notably heart failure, is expected due to demographic shifts. Such an increase places an onerous burden on healthcare infrastructures, with recurring hospital admissions and heightened mortality rates being prominent factors. Efficient chronic disease management hinges on regular ambulatory care and preemptive action. The application of intelligent computational models is showing promise as a key resource in the ongoing management of chronic diseases, particularly in forecasting disease trajectory and informing timely interventions. In this review, we explored a pioneering intelligent computational model by Diagnostic Robotics, an Israeli start-up company. This model uses data sourced from insurance claims to forecast the progression of heart failure. The goal of the model is to identify individuals at increased risk for heart failure, thus enabling interventions to be initiated early, mitigating the risk of disease worsening, and relieving the pressure on healthcare facilities, which will result in economic efficiencies.

IMAJ 2024; 26: 108–113

KEY WORDS: artificial intelligence (AI), heart failure, machine learning, predictive model, proactive management

One in six people in the United States presents with at least one chronic disease, including heart failure, diabetes, or respiratory diseases. Chronic morbidity places a substantial demand on healthcare infrastructure, incurs significant financial expenditures, and is a major cause of death. The demographic shift toward an older population in the ensuing decades is anticipated to markedly elevate the incidence of chronic diseases [1].

The trajectory of chronic diseases is critically dependent on effective ambulatory management [2,3]. Optimal proactive management is instrumental in maintaining disease equilibrium, minimizing episodes of exacerbation, and averting long-term deterioration [4]. In contrast, patients receiving suboptimal continuous care are likely to experience frequent acute exacerbations and a progressive decline in overall health compounded by co-morbidities and functional impairments. This clinical imbalance is further associated with escalated healthcare costs, predominantly stemming from recurrent emergency department visits and hospitalizations, thus imposing considerable strain on healthcare systems.

Effective allocation of ambulatory resources necessitates precise identification of patients at elevated risk of deterioration. Artificial intelligence (AI) models have emerged as a potent tool for classifying patient risk levels, due to their ability to consider many dynamic variables and embody the complex multifactorial influence they have on disease progression [5].

Diagnostic Robotics, an Israeli start-up company, is at the forefront of developing AI-based predictive models for deterioration of patients with chronic diseases. These models are operational at a juncture where proactive intervention can still significantly alter the disease course [6]. The models utilize extensive data from insurance claims, including ICD-10 diagnosis codes, current procedural terminology procedures, and National Drug Code medications. The company has developed several models pinpointing individuals at the highest risk of future deterioration. These models enable early initiation of intervention to prevent the expected worsening in their condition, thereby reducing healthcare system burdens and associated costs.

This review article will elaborate on one of the company's models, specifically designed for predicting the progression of heart failure [7].

MACHINE LEARNING: A NEW ERA IN MEDICINE

The recent decade has witnessed a remarkable evolution in AI, unlocking a spectrum of novel predictive possibilities in the medical field [8]. This growth encompasses a broad range of applications from disease risk prediction to customizing treatments based on anticipated patient responses. These advancements offer unprecedented opportunities for both researchers and clinicians, enhancing patient care quality and treatment efficacy. Machine learning, a pivotal sub-discipline within AI, leverages large-scale data sets to identify complex patterns. Such models are increasingly integral in facilitating decision-making processes within patient-provider interactions [9].

Machine learning methods can be divided into three subgroups:

- *Supervised learning:* This approach involves training models on pre-labeled datasets, where the desired outcome is predetermined. The objective is to enable these models to classify new examples not included in the initial training set.
- *Unsupervised learning:* This method autonomously uncovers patterns or relationships in data without reliance on pre-established labeling or categorization by human intermediaries. Models of this type are pivotal in identifying novel risk factors and previously unrecognized correlations among various medical conditions.
- *Reinforcement learning:* In this subgroup, models do not have access to all data upfront but rather learn through trial and error. Functioning as an *agent*, the model interacts with its environment gathering current information and executing actions aimed at maximizing rewards based on a defined policy. In healthcare, these models hold potential for refining the optimization of medication dosages or other continuous interventions, playing a key role in the advancement of personalized medicine.

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING MODELS, SUCH AS THOSE CREATED BY DIAGNOSTIC ROBOTICS, ARE KEY IN MANAGING CHRONIC DISEASES LIKE HEART FAILURE, FORECASTING PROGRESSION, AND ENABLING EARLY INTERVENTIONS TO REDUCE HEALTHCARE STRAIN.

The integration of machine learning techniques with conventional medical research can lead to the development of more precise and individualized models. These models are particularly relevant for chronic diseases like heart failure, the focus of the model discussed in this review. The showcased model utilizes supervised learning to achieve its predictive objectives.

HEART FAILURE

Chronic and acute heart failure represent principal contributors to disease burden and mortality, both in the United States and globally. Projections indicate that by 2030, approximately 8 million individuals in the United States will be afflicted with heart failure. Financial projections estimate that heart failure-related expenses will amount to around US$53.1 billion by 2030 in the United States alone [10]. The clinical progression of heart failure is characterized by repeated chronic exacerbations and a gradual decline over several years, culminating in cardiac failure. Proactive ambulatory management has proven efficacious in sustaining equilibrium in cardiac function, mitigating exacerbations, and decelerating the progression of chronic deterioration. Furthermore, proactive therapeutic strategies are instrumental in reducing emergency department visits and hospital admissions, thereby diminishing the associated economic impact and the strain on healthcare systems [2].

HEART FAILURE MODEL

Our objective was to develop an AI model that would predict preventable hospital admissions, emergency department visits, and associated medical costs in heart failure patients. We compared the AI model with traditional logistic regression modeling to evaluate the advantages of machine learning and deep learning methodologies in enhancing predictive accuracy. The model was developed using 12 years of medical information based on medical insurance claims data from a major insurance company in the United States, which was collected between January 2006 and December 2017. The information was completely anonymized and contained the entire insurance variety of the insurance company. Diagnosis codes, procedures, and medications were grouped into several accepted categories with clinical significance, such as diagnosis based on clinical classifications software (CCS) categories. Annual costs were calculated as the total insurance payments for the patient between 2016 and 2017. Extreme values from the 1st and 99th percentiles were limited by winsorization to reduce the impact of outlier observations.

The data were divided into two periods: an observation period before 1 January 2017 and a prediction period starting after 1 January 2017. During the observation period, adult patients (aged 18 and older) with heart failure who were continuously insured between 2016 and 2017 and covered by a plan that also includes drug coverage were identified. Patients with active advanced malignancy during the observation period were excluded from the sample population to avoid overestimation of preventable visit costs.

Attention was centered on three binary outcomes. The first was the occurrence of any preventable hospitalization within a 6-month period from 1 January 2017. The second outcome involved any preventable emergency department (ED) visit within the same 6-month interval based on the premise that shorter-term healthcare utilization is more clinically actionable. The third outcome assessed was any preventable costs, defined as the expenses of preventable hospitalizations and ED visits within a 1-year period starting from 1 January 2017, a time frame typically utilized in cost analysis studies.

To identify potentially preventable hospitalizations, the Agency for Healthcare Research and Quality (AHRQ) prevention quality indicators (PQI) was used [11]. PQIs define potentially preventable hospitalizations, as hospitalizations related to chronic diseases such as heart failure, diabetes, hypertension, and asthma, which could be avoided with good and timely outpatient care. This algorithm has been both validated and utilized in the past [12]. To identify potentially preventable ED visits, a combination of two validated algorithms was used. The first was an updated version of an algorithm created by Johnson et al. [13,14] that uses principal diagnosis codes to separate non-admitted ED visits into four categories: non-emergent; emergent but primary care treatable; emergent, ED care needed, but preventable; and emergent, ED care needed, and not preventable. A second procedure-based algorithm created by Ballard and co-authors [15] was then used on the remaining unclassified visits to capture additional preventable visits where there was an absence of ED-indicating procedures.

ARTIFICIAL INTELLIGENCE ENABLES EARLY INTERVENTIONS, THUS REDUCING HEALTHCARE STRAIN.

Patient features or predictors were classified into two primary categories [16]. The first category comprised knowledge-driven features, manually compiled by domain experts. This set, consisting of 939 features, encapsulated demographic data and medical background information of the patients such as episode counts and trends, hospital length of stay, readmission rates, costs, co-morbidity indicators, major procedure indicators, and chronic medications. Specific heart failure-related features included clinical subtypes, episodes, procedures, and heart failure medication adherence indicators. The second category encompassed data-driven features, which were machine learning-based representations of each patient's medical codes. They served as inputs for machine learning predictive models. These features were generated using the word2vec algorithm, a natural language processing method that creates a feature vector (i.e., an array of numbers) for each medical code in a patient's history. To form vectors representing each patient, rather than a single code, the sets of vectors for patients were summed. The results were represented in two ways: as single patient-level vectors (non-sequential vector inputs), containing the sum of a patient's 11-year medical history and as temporal patient-level vectors (sequential vector inputs) comprising 36 consecutive monthly vectors with each vector summing the medical codes for one month.

MACHINE LEARNING IN HEALTHCARE ALLOWS FOR THE IDENTIFICATION OF COMPLEX PATTERNS IN LARGE DATA SETS, LEADING TO MORE ACCURATE RISK CLASSIFICATION AND PROACTIVE DISEASE MANAGEMENT.

Initially, the final sample was randomly divided into three groups: training, validation, and testing, in a 7:2:1 ratio. The training sample was used to develop the model, while the validation and testing samples were used for model tuning and results presentation, respectively. Using this data, five predictive models were created for comparison: two conventional logistic regression (LR) models with a limited feature set, a more comprehensive LR model incorporating the full array of knowledge-driven features, and two machine learning models employing varied methodologies. The first traditional LR model encompassed basic attributes like age, sex, and disease risk scores (CCS score [17] and chronic condition indicator score [18]). The second traditional LR model expanded these scores by incorporating cost-related features (inpatient, outpatient specialists, pharmacy, and primary care costs). These two models have been widely utilized in the United States for creating risk scores in diagnosis-based and pharmacy-based cost-prediction tools [19,20]. The enhanced LR model also used the complete set of the 939 knowledge-driven features. The two machine learning models were regarded as sequential or non-sequential, according to the models' input features. Machine learning models using non-sequential inputs (i.e., single patient-level vectors) included feedforward neural network and gradient boosting model. Deep learning models using sequential inputs (i.e. temporal patient-level vectors) included convolutional neural networks and long-short term memory with an attention mechanism. For each of the approaches the best performing models were chosen based on the evaluation metrics [21,22].

The analysis of the results initially focused on comparing patient features across the three groups (training, validation, testing) using chi-square tests and analysis of variance (ANOVA), with statistical significance considered at $P < 0.05$. The second phase involved evaluating the preventable hospitalizations and ED visits using the precision at K method. This method considers the patients ranked highest by the model in the top K percent (ranging from 1% to 10%), where precision is the positive predictive value. The analysis of preventable costs was performed using the cost capture metric, defined as the ratio between the predicted preventable costs and the actual preventable costs. This analysis considered the costs in the top K%.

RESULTS

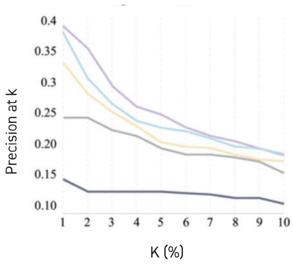
Of the five candidate modeling approaches evaluated, the sequential deep learning models consistently provided the best predictive performance across all three outcomes and were closely followed by the non-sequential machine learning models [Figure 1].

The models also revealed the strongest predictors for each of the three outcomes: ED visits, preventable hospitalizations, and potential cost savings. This calculation was figured by taking predefined predictive variables to compare patients and calculate each patient's relative risk. Each of these variables was assigned a score between 0 to 100 according to its relative impact on the patient's risk. Figure 2 shows the 15 variables found to be the most influential on each of the model's three outcomes: preventable hospitalizations, ED visits, and potential cost savings.

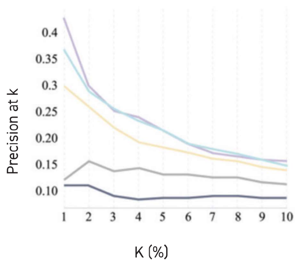
Figure 1. Models' precision and cost capture at K%

CNN = convolutional neural networks, FNN = feedforward neural network, GBM = gradient boosting model, LR = logistic regression, LSTM = long-short term memory

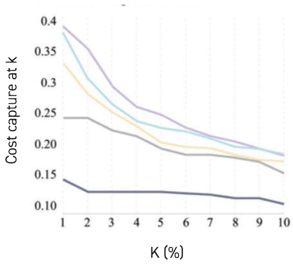
[A] Preventable hospitalizations



[B] Preventable emergency department visits



[C] Preventable cost



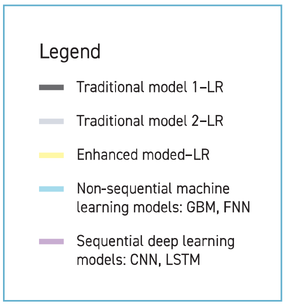
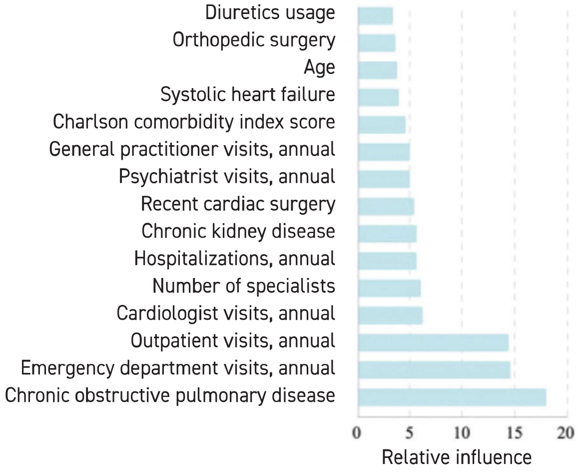
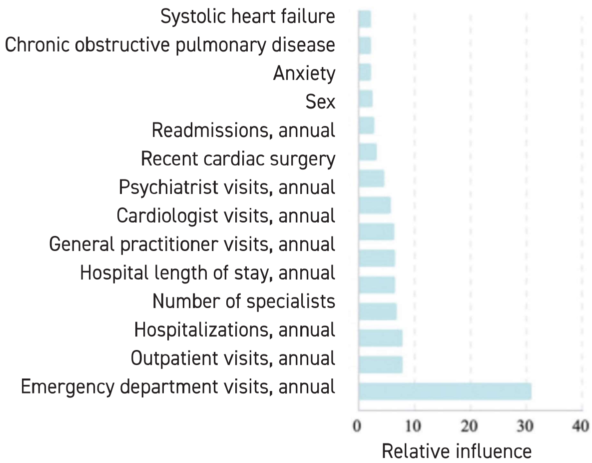


Figure 2. Top 15 predictors for preventable hospitalizations, emergency department visits, and high costs

[A] Top 15 predictors for preventable hospitalizations



[B] Top 15 predictors for preventable emergency department visits



[C] Top 15 predictors for preventable cost

In the United States, insurance companies actively reach out to patients with heart failure after hospitalization and to patients with high treatment costs due to the disease. These patients are recruited to join various intervention programs designed to improve their health, and thus, prevent further deteriorations leading to hospitalization. The patients were divided into two groups. The first group consisted of patients flagged by the insurance company's standard urgency predictor, and the second comprised patients our model identified as high risk. Notably, only 10% of the congestive heart failure patients flagged by our model were also identified by the insurance company's health plan. Each of these groups was then paired with a corresponding control group, which was not proactively engaged despite being flagged by the systems. In a difference-in-difference experiment, it was found that our model had the potential to lead to a 1.9-fold reduction in hospitalizations compared to the traditional system of the insurance company.

DISCUSSION

The application of AI and machine learning models enables identification of patients at heightened risk of disease progression. These models can enable healthcare providers to intervene earlier, alter the disease's trajectory, and reduce emergency department admissions and hospitalizations. This proactive approach not only enhances patient outcomes but also alleviates the burden on healthcare systems and reduces costs, a crucial aspect considering the escalating prevalence of chronic diseases.

However, the adoption of such AI models in healthcare raises several considerations. The complexity and often opaque nature of these models, commonly referred to as the *black box*, where sometimes only the input and output are understandable, present challenges in terms of comprehending and trusting their decision-making processes. This lack of transparency can be a barrier to the wider acceptance and integration of AI in clinical settings. Addressing these concerns is crucial for the future development and implementation of AI in healthcare.

CONCLUSIONS

As AI continues to evolve, it will play an increasingly vital role in healthcare by offering more accurate predictions, personalized treatment plans, and improved patient outcomes.

Correspondence
Dr. A. Peri
Diagnostic Robotics, Tel Aviv 6473926, Israel
Email: almap@diagnosticrobotics.com; alma.peri@gmail.com

References

1. Ansah JP, Chiu CT. Projecting the chronic disease burden among the adult population in the United States using a multi-state population model. *Front Public Health* 2023; 10: 1082183.
2. Figueroa JF, Joynt Maddox KE, Beaulieu N, Wild RC, Jha AK. Concentration of potentially preventable spending among high-cost Medicare subpopulations: an observational study. *Ann Intern Med* 2017; 167 (10): 706-13.
3. Joynt KE, Gawande AA, Orav EJ, Jha AK. Contribution of preventable acute care spending to total spending for high-cost Medicare patients. *JAMA* 2013; 309 (24): 2572-8.
4. Heidenreich PA, Bozkurt B, Aguilar D, et al. 2022 AHA/ACC/HFSA Guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* 2022; 145 (18): e895-1032.
5. Javaid M, Haleem A, Pratap Singh R, Suman R, Rab S. Significance of machine learning in healthcare: features, pillars and applications. *Int J Intell Netw* 2022; 3: 58-73.
6. Diagnostic Robotics. Transforming Healthcare with AI: Empowering Health Plans and Providers with Precise, Predictive Analytics. [Available from https://www.diagnosticrobotics.com/]. [Accessed 30 June 2023].
7. Lewis M, Elad G, Beladev M, et al. Comparison of deep learning with traditional models to predict preventable acute care use and spending among heart failure patients. *Sci Rep* 2021; 11 (1): 1164.
8. Obermeyer Z, Emanuel EJ. Predicting the future–big data, machine learning, and clinical medicine. *N Engl J Med* 2016; 375 (13): 1216-9.
9. Deo RC. Machine learning in medicine. *Circulation* 2015; 132 (20): 1920-30.
10. Ziaeian B, Fonarow GC. Epidemiology and aetiology of heart failure. *Nat Rev Cardiol* 2016; 13 (6): 368-78.
11. Farquhar M. AHRQ Quality Indicators. In: Hughes RG, editor. Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008 Apr. Chapter 45. PMID: 21328764.
12. Jiang HJ, Russo CA, Barrett ML. Nationwide Frequency and Costs of Potentially Preventable Hospitalizations, 2006. 2009 Apr. In: Healthcare Cost and Utilization Project (HCUP) Statistical Briefs [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2006 Feb–. Statistical Brief #72. PMID: 21510033.
13. Johnston KJ, Allen L, Melanson TA, Pitts SR. A “Patch” to the NYU emergency department visit algorithm. *Health Serv Res* 2017; 52 (4): 1264-76.
14. Faculty & Research | NYU Wagner [Internet]. [Available from https://wagner.nyu.edu/faculty/billings/nyued-background]. [Accessed 1 July 2023].
15. Ballard DW, Price M, Fung V, et al. Validation of an algorithm for categorizing the severity of hospital emergency department visits. *Med Care* 2010; 48 (1): 58-63.
16. Billings J, Parikh N, Mijanovich T. Emergency department use: the New York Story. *Issue Brief (Commonw Fund)* 2000; (434): 1-12.
17. Johnston KJ, Allen L, Melanson TA, Pitts SR. A "Patch" to the NYU Emergency Department Visit Algorithm. *Health Serv Res* 2017; 52 (4): 1264-76.

18. Ballard DW, Price M, Fung V, et al. Validation of an algorithm for categorizing the severity of hospital emergency department visits. *Med Care* 2010; 48 (1): 58-63.
19. Fleishman JA, Cohen JW. Using information on clinical conditions to predict high-cost patients. *Health Serv Res* 2010; 45: 532-52.
20. Centers for Medicare and Medicaid Services. March 31, 2016, HHS-Operated Risk Adjustment Methodology Meeting. [Available from https://www.cms.gov/cciio/resources/forms-reports-and-other-resources/downloads/ra-march-31-white-paper-032416.pdf].
21. Cumming RB, Knutson D, Cameron BA, Brian D. A Comparative Analysis of Claims-Based Methods of Health Risk Assessment for Commercial Populations. Final report to the Society of Actuaries. 2002 May 24. [Available from https://www.soa.org/Files/Research/Projects/2005-comp-analysis-methods-commercial-populations.pdf].
22. Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE J Biomed Heal Inform* 2018; 22: 1589-604.

Capsule

Mechanosensing by T cells

Tissue-resident CD8+ T (TRM) cells are constantly surveilling organs and tissues for the presence of uninvited microbes. Previous studies showed that TRM cell migration is triggered by chemoattractant and adhesion molecule signaling, which facilitates the rapid detection of infected cells. More recent evidence indicated that TRM cells within submandibular salivary glands display different motility patterns exclusive of chemosensing. **Ruef** and co-authors showed that submandibular salivary gland TRM cells from virally infected mice display spontaneous retrograde F-actin flow as a means of force-generated translocation. Similar patterns of locomotion were detected in TRM cells from other exocrine glands and were dependent on the sensing of changes in mechanical loads through signals triggered by nuclear deformation.

Sci Immunol 2023; 8 (90): eadd5724

Eitan Israeli

Capsule

Prosopagnosia: face blindness and its association with neurological disorders

Josephs and **Josephs** assessed demographic, clinical, and imaging characteristics, and neurological and neuropathological disorders associated with a diagnosis of prosopagnosia (loss of facial recognition) in a large cohort. Patients were categorized as developmental versus acquired. Those with acquired prosopagnosia were further subdivided into degenerative versus non-degenerative, based on neurological etiology. The authors assessed regional involvement on 18F-fluorodeoxyglucose PET and MRI of the right and left frontal, temporal, parietal, and occipital lobes. The Intake and Referral Center at the Mayo Clinic identified 487 patients with possible prosopagnosia, of which 336 met study criteria for probable or definite prosopagnosia. Ten patients, 80.0% male, had developmental prosopagnosia including one with Niemann-Pick type C, and another with a Forkhead-box G1 gene mutation. Of the 326 with acquired prosopagnosia, 235 (72.1%) were categorized as degenerative, 91 (27.9%) as non-degenerative. The most common degenerative diagnoses were posterior cortical atrophy, primary prosopagnosia syndrome, Alzheimer's disease dementia, and semantic dementia, with each diagnosis accounting for > 10% of this group. The most common non-degenerative diagnoses were infarcts (ischemic and hemorrhagic), epilepsy-related, and primary brain tumors, each accounting for > 10%. The authors identified a group of patients with non-degenerative transient prosopagnosia in which facial-recognition loss improved or resolved over time. These patients had migraine-related prosopagnosia, posterior reversible encephalopathy syndrome, delirium, hypoxic encephalopathy, and ischemic infarcts. On 18F-fluorodeoxyglucose PET, the temporal lobes proved to be the most frequently affected regions in 117 patients with degenerative prosopagnosia, while in 82 patients with non-degenerative prosopagnosia MRI revealed the right temporal and right occipital lobes as most affected by a focal lesion. The most common pathological findings in those with degenerative prosopagnosia were frontotemporal lobar degeneration with hippocampal sclerosis, and mixed Alzheimer's and Lewy body disease pathology. In this large case series of patients diagnosed with prosopagnosia, they observed that facial-recognition loss occurs across a wide range of acquired degenerative and non-degenerative neurological disorders, most commonly in males with developmental prosopagnosia. The right temporal and occipital lobes, and connecting fusiform gyrus, are key areas.

Brain Commun 2024; https://doi.org/10.1093/braincomms/fcae002

Eitan Israeli