

Diagnosing Mental Health Disorders Through AI Facial Expression Evaluation

Originally published at <https://www.unite.ai/diagnosing-mental-disorders-through-ai-facial-expression-evaluation/>



Researchers from Germany have developed a method for identifying mental disorders based on facial expressions interpreted by computer vision.

The new approach can not only distinguish between unaffected and affected subjects, but can also correctly distinguish depression from schizophrenia, as well as the degree to which the patient is currently affected by the disease.

The researchers have provided a composite image that represents the control group for their tests (on the left in the image below) and the patients who are suffering from mental disorders (right). The identities of multiple people are blended in the representations, and neither image depicts a particular individual:



Source: <https://arxiv.org/pdf/2208.01369.pdf>

Individuals with affective disorders tend to have raised eyebrows, leaden gazes, swollen faces and hang-dog mouth expressions. To protect patient privacy, these composite images are the only ones made available in support of the new work.

Until now, facial affect recognition has been primarily used as a potential tool for basic diagnosis. The new approach, instead, offers a possible method to evaluate patient progress throughout treatment, or else (potentially, though the paper does not suggest it) in their own domestic environment for outpatient monitoring.

The paper states*:

'Going beyond machine diagnosis of depression in affective computing, which has been developed in [previous studies](#), we show that the measurable affective state estimated by means of computer vision contains far more information than the pure categorical classification.'

The researchers have dubbed this technique *Opto Electronic Encephalography* (OEG), a completely passive method of inferring mental state by facial image analysis instead of topical sensors or ray-based medical imaging technologies.

The authors conclude that OEG could potentially be not just a mere secondary aide to diagnosis and treatment, but, in the long term, a potential replacement for certain evaluative parts of the treatment pipeline, and one that could cut down on the time necessary for patient monitoring and initial diagnosis. They note:

'Overall, the results predicted by the machine show better correlations compared to the pure clinical observer rating based questionnaires and are also objective. The relatively short measurement period of a few minutes for the computer vision approaches is also noteworthy, whereas hours are sometimes required for the clinical interviews.'

However, the authors are keen to emphasize that patient care in this field is a multi-modal pursuit, with many other indicators of patient state to be considered than just their facial expressions, and that it is too early to consider that such a system could entirely substitute traditional approaches to mental disorders. Nonetheless, they consider OEG a promising adjunct technology, particularly as a method to grade the effects of pharmaceutical treatment in a patient's prescribed regime.

The [paper](#) is titled *The Face of Affective Disorders*, and comes from eight researchers across a broad range of institutions from the private and public medical research sector.

Data

(The new paper deals mostly with the various theories and methods that are currently popular in patient diagnosis of mental disorders, with less attention than is usual to the actual technologies and processes used in the tests and various experiments)

Data-gathering took place at University Hospital at Aachen, with 100 gender-balanced patients and a control group of 50 non-affected people. The patients included 35 sufferers from schizophrenia and 65 people suffering from depression.

For the patient portion of the test group, initial measurements were taken at the time of first hospitalization, and the second prior to their discharge from hospital, spanning an average interval of 12 weeks. The control group participants were recruited arbitrarily from the local population, with their own induction and 'discharge' mirroring that of the actual patients.

In effect, the most important 'ground truth' for such an experiment must be diagnoses obtained by approved and standard methods, and this was the case for the OEG trials.

However, the data-gathering stage obtained additional data more suited for machine interpretation: interviews averaging 90 minutes were captured over three phases with a Logitech c270 consumer webcam running at 25fps.

The first session comprised of a standard [Hamilton interview](#) (based on research originated around 1960), such as would normally be given on admission. In the second phase, unusually, the patients (and their counterparts in the control group) were shown videos of a series of facial expressions, and asked to mimic each of these, while stating their own estimation of their mental condition at that time, including emotional state and intensity. This phase lasted around ten minutes.

In the third and final phase, the participants were shown 96 videos of actors, lasting just over ten seconds each, apparently recounting intense emotional experiences. The participants were then asked to evaluate the emotion and intensity represented in the videos, as well as their own corresponding feelings. This phase lasted around 15 minutes.

Method

To arrive at the mean average of the captured faces (see first image, above), emotional landmarks were captured with the [EmoNet](#) framework. Subsequently, correspondence between the face shape and the mean (averaged) face shape was determined through [piecewise affine transformation](#).

[Dimensional emotion recognition](#) and [eye gaze prediction](#) was carried out on each landmark segment identified in the previous stage.

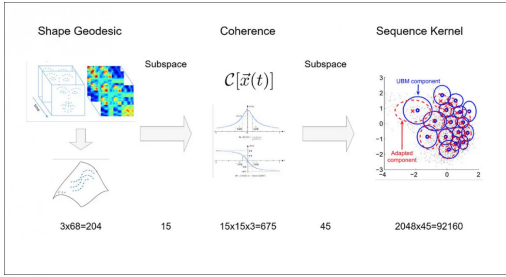
At this point, audio-based emotion inference has indicated that a teachable moment has arrived in the patient's mental state, and the task is to capture the corresponding facial image and develop that dimension and domain of their affect state.

Automatic emotion analysis from faces in-the-wild



(In the video above, we see the work developed by the authors of the dimensional emotion recognition technologies used by the researchers for the new work).

The shape geodesic of the material was computed for each frame of the data, and Singular Value Decomposition ([SVD](#)) reduction applied. The resultant time series data was eventually modeled as a [VAR](#) process, and then further reduced via SVD prior to [MAP adaptation](#).



Workflow for the geodesic reduction process.

The valence and arousal values in the EmoNet network were also similarly processed with VAR modelling and sequence kernel computation.

Experiments

As explained earlier, the new work is primarily a medical research paper rather than a standard computer vision submission, and we refer the reader to the paper itself for in-depth coverage of the diverse OEG experiments run by the researchers.

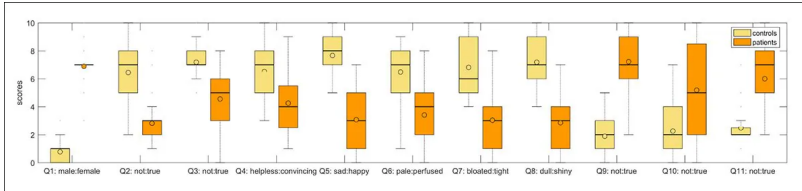
Nonetheless, to summarize a selection of them:

Affective Disorder Cues

Here 40 participants (not from the control or patient group) were asked to rate the evaluated mean faces (see above) in respect to a number of questions, without being informed of the context of the data. The questions were:

- What is the gender of the two faces?
- Do the faces have an attractive appearance?
- Are these faces trustworthy persons?
- How do you assess the ability of these people to act?
- What is the emotion of the two faces?
- What is the skin appearance of the two faces?
- What is the impression of the gaze?
- Do the two faces have droopy mouth corners?
- Do the two faces have raised eye brows?
- Are these persons clinical patients?

The researchers found that these blind evaluations correlated to the registered state of the processed data:



Box plot results for the 'mean face' survey.

Clinical Assessment

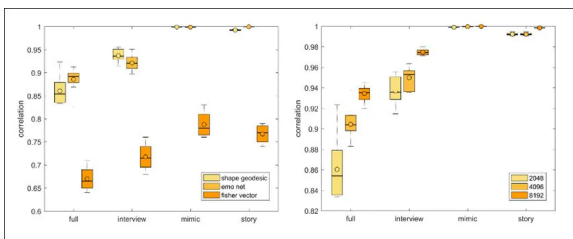
To gauge the utility of OEG in initial assessment, the researchers first evaluated how effective standard clinical assessment is by itself, measuring levels of improvement between the induction and the second phase (by which time the patient is typically receiving drug-based treatments).

The researchers concluded that status and symptom severity could be well-assessed by this method, achieving a correlation of 0.82. However, an accurate diagnosis of either schizophrenia or depression proved more challenging, with the standard method only obtaining a score of -0.03 at this early stage.

The authors comment:

'In essence, the patient status can be determined relatively well using the usual questionnaires. However, that is essentially all that can be concluded from it. Whether someone is depressed or rather schizophrenic is not indicated. The same applies to the treatment response.'

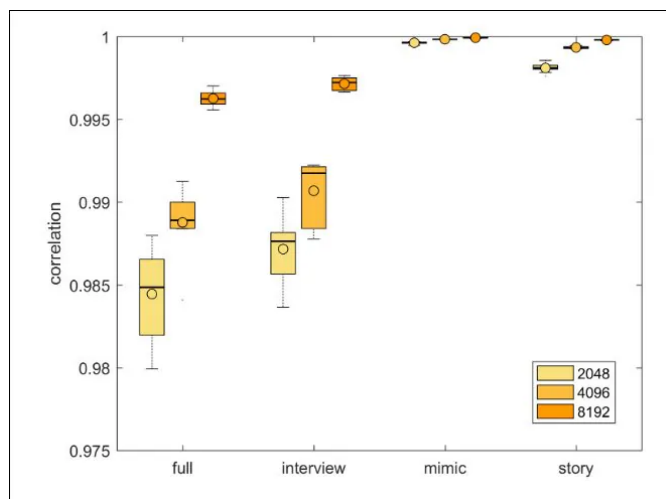
The results from the machine process were able to obtain higher scores in this problem area, and comparable scores for the initial patient evaluation aspect:



Higher numbers are better. On the left, standard interview-based evaluation accuracy results across four phases of the testing architecture; on the right, machine-based results.

Disorder Diagnosis

Distinguishing depression from schizophrenia via static face images is not a trivial matter. Cross-validated, the machine process was able to obtain high accuracy scores across the various phases of the trials:



In other experiments, the researchers were able to demonstrate evidence that OEG can perceive patient improvement through pharmacological treatment, and general treatment of the disorder:

'The causal inference over the empirical prior knowledge of the data collection adjusted the pharmacological treatment in order observe a return to the physiological regulation of the facial dynamics. Such a return couldn't be observed during the clinical prescription.'

'At the moment it is not clear whether such a machine based recommendation would indeed result to a significant better success of therapy. Especially because it is known which side effects medications can have over a long period of time.'

'However, [these kinds] of patient-tailored approaches would break the barriers of the common categorical classification schematic still dominantly used in daily life.'

* My conversion of the authors' inline citations to hyperlinks.

First published 3rd August 2022.

Writer on machine learning, domain specialist in human image synthesis. Former head of research content at Metaphysic.ai, until its dissolution into DNEG's Brahma.ai.

Portfolio site: martinanderson.ai

Contact: martin@martinanderson.ai



Discover More