

EXTRAPOLER DU MACAQUE À L'HUMAIN ?

Les travaux de Doris Tsao et son équipe constituent une prouesse technique dont seuls sont capables quelques laboratoires au monde. Se pose toutefois la question de leur validité pour élucider les mécanismes neuronaux de reconnaissance des visages chez l'homme.

Celui-ci se fonde essentiellement sur le visage pour reconnaître des individus et faire la différence entre des visages connus et inconnus. Alors qu'un adulte est capable de reconnaître plusieurs milliers de visages rapidement et

automatiquement, le singe rhésus ou le macaque a besoin de centaines d'essais d'entraînement avec les mêmes images pour atteindre des performances modestes. De plus, contrairement à l'homme, les performances du macaque sont identiques pour un visage présenté à l'endroit ou à l'envers et ne dépendent pas du degré de familiarité du visage (voir l'article de B. Rossion et J. Taubert cité en bibliographie).

Cela n'est guère surprenant : l'homme et le macaque diffèrent par 25-30 millions d'années d'évolution, et le cerveau humain compte 16 fois plus de neurones. Surtout, le macaque ne possède pas les structures cérébrales clés pour la reconnaissance du visage chez l'homme, en particulier le gyrus fusiforme du cortex occipitotemporal ventral, et une dominance de l'hémisphère droit. Pourquoi, dès lors, trouve-t-on des régions qui répondent

spécifiquement aux visages dans la partie latérale du lobe temporal du macaque et parvient-on à y décoder des « identités faciales » ? D'une part, parce que cette espèce décode dans le visage de congénères de simples expressions faciales, l'orientation de la tête et du regard. D'autre part, parce que le bombardement d'images de visages (humains) à discriminer dans des conditions artificielles de laboratoire entraîne, ou augmente, l'activité de régions cérébrales pour ces catégories d'images dans le cerveau du macaque. Il n'est alors pas surprenant que des images physiques différentes activent des réponses de populations de neurones différentes dans ces régions, permettant un décodage fiable de leur « identité ».

Les capacités de reconnaissance du visage chez l'homme vont cependant bien au-delà de la discrimination d'images 2D ; elles se fondent sur des

propriétés physiques, mais également sémantiques et contextuelles. Cela permet de reconnaître un visage familier malgré les différences importantes entre les diverses vues de ce visage, celle d'une caricature par exemple.

Malgré l'intérêt des travaux sur les singes, il nous semble que la compréhension des mécanismes de reconnaissance des visages chez l'humain passera essentiellement par l'étude de notre espèce, par exemple de patients cérébrolésés ayant perdu la capacité de reconnaissance faciale (prosopagnosie), la neuroimagerie et les enregistrements intracérébraux.

**BRUNO ROSSION,
JACQUES JONAS
ET LAURENT KOESSLER**
CNRS (CRAN, UMR7039),
université de Lorraine et service
de neurologie du CHRU de Nancy

Pour générer ces caractères de forme et d'aspect, nous sommes partis d'une grande base de données de photos de visages. Pour chaque visage, nous avons placé un ensemble de marqueurs sur les traits principaux. La position de ces marqueurs décrit la forme du visage. À partir de ces formes variées, nous avons calculé un visage moyen. Nous avons ensuite transformé chaque photo de la base de données afin que ses principaux traits correspondent exactement à ceux du visage moyen.

Les images résultantes constituent l'aspect des visages indépendamment de la forme. Nous avons ensuite effectué une « analyse en composantes principales », sur les caractères de forme et sur les caractères d'aspect, pour l'ensemble des visages. Il s'agit d'une technique mathématique qui détermine les axes, ou dimensions, décrivant le mieux la variabilité au sein d'un ensemble complexe de données.

En prenant 25 composantes principales pour la forme et 25 composantes principales pour l'aspect, nous avons créé un espace facial à 50 dimensions (voir l'encadré page 49). Chaque visage est représenté dans cet espace par un point (ou vecteur) ayant 50 coordonnées (ou composantes), alors que dans l'espace usuel, chaque point représente une position et est repéré par 3 coordonnées.

Dans notre expérience, nous avons pris au hasard 2000 visages et les avons présentés à un singe tout en enregistrant les neurones de deux

zones faciales. Nous avons constaté que presque chaque cellule présentait des réponses dépendant linéairement d'un sous-ensemble des 50 caractères, ce qui est cohérent avec mes expériences précédentes sur les visages dessinés. Mais nous avions une meilleure idée de la raison pour laquelle cela est important. Si une cellule faciale présente une réponse linéaire à différents traits, on peut approximer sa réponse par une simple somme pondérée des traits faciaux, les poids de cette somme caractérisant la dépendance linéaire de la réponse. En termes plus précis, les réponses des neurones sont données par la relation :

Réponses des cellules faciales = $(M) \times (50 \text{ traits faciaux})$, où M est la matrice de pondération, tableau de nombres où chaque ligne correspond aux poids associés à une cellule faciale. On peut alors simplement inverser cette équation pour prédire le visage montré au sujet à partir des réponses des cellules faciales : $(50 \text{ traits faciaux}) = (M^{-1}) \times (\text{réponses des cellules faciales})$.

Pour tester cette relation, nous avons utilisé les réponses à tous les 2000 visages sauf un afin de déterminer la matrice de pondération, puis nous avons calculé les 50 traits du visage exclu à partir des réponses des cellules faciales et de notre formule. Étonnamment, la prédiction s'est révélée presque indistinguishable du visage réel.

Lors d'une conférence tenue en 2015 à Ascona, en Suisse, j'ai présenté nos ➤

➤ découvertes sur la reconstitution de visages à partir de l'activité neuronale. Après mon exposé, Rodrigo Quian Quiroga, actuellement à l'université de Leicester, en Angleterre, qui avait découvert en 2005 le célèbre «neurone Jennifer Aniston» (un neurone unique qui s'active à la vue de photos de l'actrice Jennifer Aniston) dans le lobe temporal médian, m'a posé une question. Il m'a demandé comment je faisais le lien entre mes cellules et son idée selon laquelle des neurones individuels réagissent au visage de personnes particulières, comme le «neurone Jennifer Aniston».

UN PARI GAGNANT-GAGNANT

Je lui ai répondu que je pensais que nos cellules pourraient être les éléments de base de ses cellules, sans trop réfléchir à la façon dont cela fonctionnerait. Cette nuit-là, privée de sommeil à cause du décalage horaire, j'ai pris conscience d'une différence majeure entre nos cellules faciales et les siennes. J'avais décrit dans mon exposé que nos cellules faciales calculent leur réponse comme des sommes pondérées de différentes caractéristiques faciales. Or ce calcul revient à une opération mathématique nommée produit scalaire, dont la représentation géométrique est la projection d'un vecteur sur un axe (à l'instar de l'ombre d'un piquet projetée sur le sol par la lumière du soleil).

Grâce à mes souvenirs d'algèbre linéaire, je me suis rendu compte qu'il était alors possible de construire pour chaque cellule un vaste ensemble de visages distincts tous représentés dans l'espace à 50 dimensions par des vecteurs perpendiculaires à l'axe de projection, donc de projection nulle. Ainsi, tous ces visages activeraient la cellule exactement de la même façon.

Et cela suggérerait que les cellules des zones faciales sont fondamentalement différentes des cellules envisagées par Rodrigo. Cela démolissait la vague intuition que tout le monde partageait au sujet des cellules faciales, à savoir qu'elles réagiraient à des visages particuliers.

Au petit-déjeuner du lendemain matin, j'espérais trouver Rodrigo pour lui en parler. Étonnamment, quand il est arrivé, il m'a dit avoir eu exactement la même idée. Rodrigo m'a alors proposé une sorte de pari gagnant-gagnant. Si une cellule donnait effectivement la même réponse à différents visages, j'enverrais une bonne bouteille de vin à Rodrigo. Si, au contraire, la prédiction ne se vérifiait pas, il m'enverrait du vin comme prix de consolation.

De retour à notre laboratoire à Caltech, Le Chang a d'abord déterminé l'axe privilégié par un neurone donné (l'axe de projection évoqué plus haut) à partir de ses réponses aux 2000 visages. Il a ensuite généré une gamme de visages tous représentés par des vecteurs perpendiculaires à l'axe privilégié. De façon

remarquable, tous ces visages ont suscité exactement la même réponse du neurone. La semaine suivante, Rodrigo a reçu une excellente bouteille de cabernet. Cette découverte prouvait que les cellules faciales du cortex IT ne codent pas l'identité de personnes particulières. Au lieu de cela, elles effectuent une projection vectorielle sur un axe, un calcul beaucoup plus abstrait.

Il semblerait alors que le neurone Jennifer Aniston n'existe pas, du moins pas dans le cortex IT. Mais des neurones individuels répondant sélectivement à des individus familiers pourraient être à l'œuvre dans une partie du cerveau qui traite le signal de sortie des cellules faciales. Les régions de stockage de la mémoire, l'hippocampe et les zones adjacentes, pourraient contenir des cellules permettant de reconnaître un individu.

La reconnaissance faciale dans le cortex IT repose donc sur un ensemble d'environ 50 nombres représentant les mesures d'un visage le long d'un ensemble de 50 axes d'un espace abstrait. Et la découverte de ce code extrêmement simple pour l'identification des visages a des implications majeures pour notre compréhension de la représentation visuelle des objets en général. Il est possible que tout le cortex IT soit organisé selon le même principe que les zones faciales, des groupes de neurones codant différents ensembles d'axes afin de représenter un objet. Nous menons actuellement des expériences pour tester cette hypothèse.

PIERRE DE ROSETTE NEURONALE

Si vous vous rendez au British Museum, à Londres, vous verrez un objet étonnant, la pierre de Rosette, sur laquelle le même décret de Memphis est gravé en trois langues différentes : en hiéroglyphes égyptiens, en démotique et en grec ancien. C'est grâce à cette pierre de Rosette, parce que les philologues connaissaient le grec ancien, qu'ils ont pu déchiffrer les hiéroglyphes égyptiens et démotiques.

De même, les visages, les zones faciales et le cortex IT constituent une pierre de Rosette neuronale, mais qui est encore en cours de déchiffrement. En montrant des photos de visages à des singes, nous avons découvert des zones faciales et appris comment leurs neurones détectent et identifient les visages. À son tour, la compréhension des principes de codage dans le réseau des zones faciales aidera peut-être à comprendre l'organisation de tout le cortex IT et à révéler ainsi comment le cerveau assure cette fonction essentielle consistant à identifier les objets, qu'il s'agisse d'un visage, d'un animal, d'une chaise ou d'un caillou. Peut-être même comprendra-t-on aussi comment les objets *imaginés* – à l'instar des courbes qui m'avaient intriguée lorsque j'étais au lycée – sont codés par le cerveau. ■

BIBLIOGRAPHIE

B. Rossion et J. Taubert, **What can we learn about human individual face recognition from experimental studies in monkeys ?**, *Vision Research*, vol. 157, pp. 142-158, 2019.

L. Chang et D. Y. Tsao, **The code for facial identity in the primate brain**, *Cell*, vol. 169(6), pp. 1013-1028, 2017.

R. Quian Quiroga, **How do we recognize a face ?**, *Cell*, vol. 169(6), pp. 975-977, 2017.