

Brain Evaluation: Failure Cases

In this survey you will review cases that were rejected by our automatic verifier model. The model assigns each rejection to one of seven error types. Whenever an essential item is involved, we provide a brief explanation.

For each MCQ you will see:

- the image(s)
- the original caption from the paper
- the paragraph in which the figure was cited ("Context")
- the generated MCQ and its proposed answer
- the essential item that triggered rejection and the verifier's explanation

Please keep in mind that the verifier model is imperfect and may be wrong. We ask you to judge whether you agree with the model's decision for each case and to provide your comments.

Because some context can be very long, we recommend you to skip this part for each question. You can refer to context part when necessary.

1. Caption: Brain maps of the voxel-wise ICCs of manually matching brain network using two-stage DBN.

Context: <none>

Q: Which brain network shows the highest average inter-rater reliability (ICC) across all three layers?

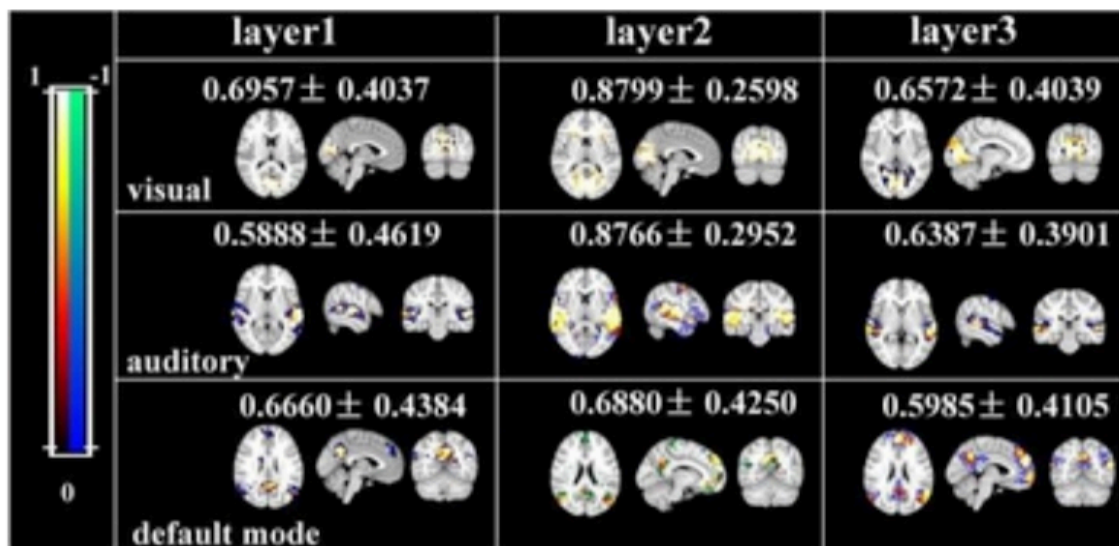
Options:

- A. Visual network
- B. Auditory network
- C. Default mode network
- D. Layer 1
- E. Layer 3

Answer: A

Verifier response: **Option Type Consistency** NOT PASS: 'Layer 1' and 'Layer 3' are not network types.

"**Option Type Consistency**" means "All options are of the same semantic type".



Mark only one oval.

- ☐ Yes, I agree with verifier's idea.
- ☐ Other: _____

2. Caption: MRI Head—Subsequent MRI imaging of the brain which showed progression of the mass with associated bony lysis.

Context: In September the MRI scan (see <xref rid="omaa075f2" ref-type="fig">Fig. 2</xref>) showed the lesion had grown significantly. It now extended across the cribriform plate, into the nasal cavity and it demonstrated associated bony lysis. These aggressive changes made the diagnosis of meningioma unlikely; as such, a transsphenoidal biopsy was obtained.) showed the lesion had grown significantly. It now extended across the cribriform plate, into the nasal cavity and it demonstrated associated bony lysis. These aggressive changes made the diagnosis of meningioma unlikely; as such, a transsphenoidal biopsy was obtained.

Q: Which finding is most consistent with the MRI image?

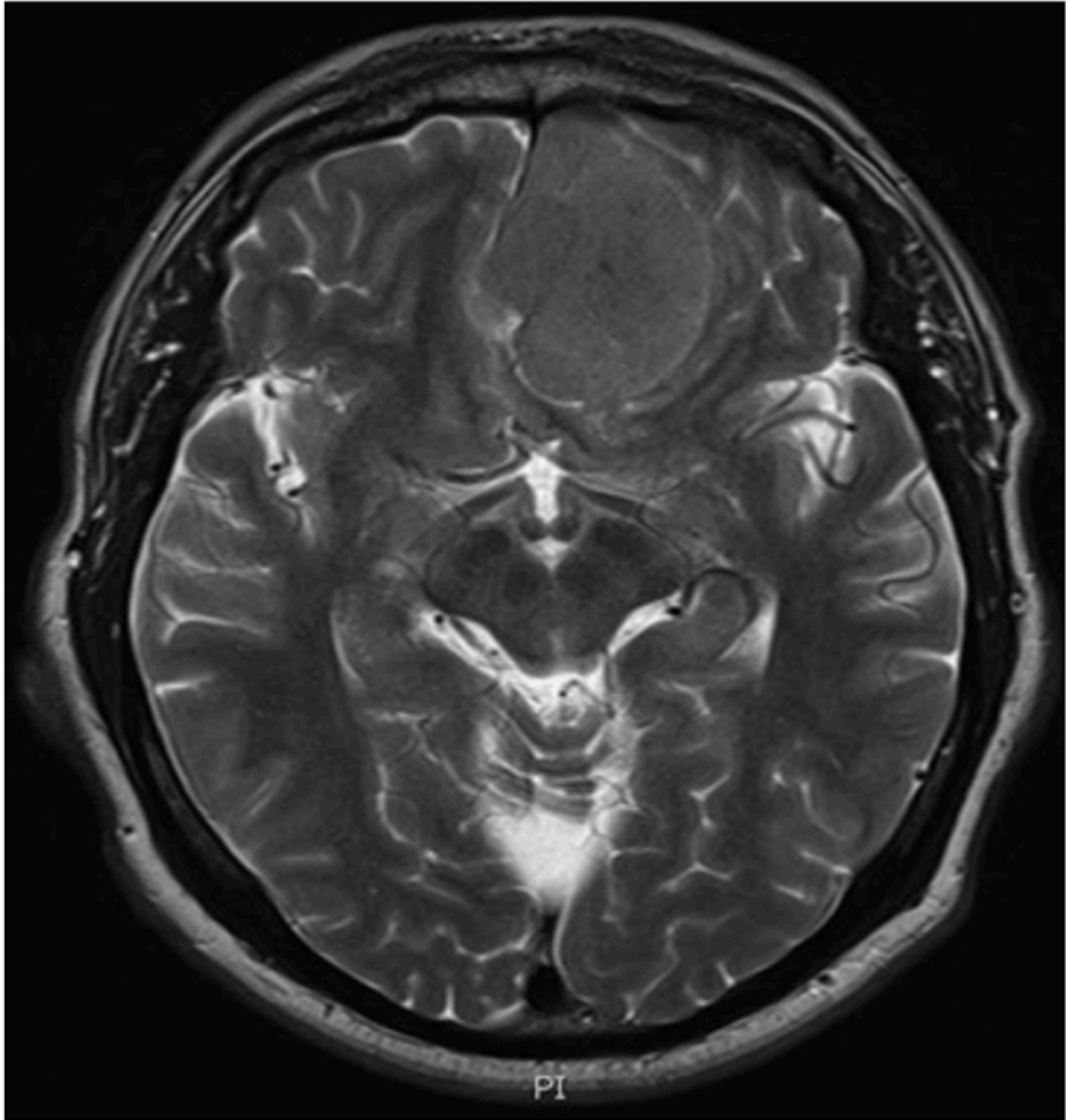
Options:

- A. A well-circumscribed mass with homogeneous enhancement
- B. A mass with associated bony lysis and extension into adjacent structures
- C. Diffuse cerebral atrophy with ventricular enlargement
- D. Multiple small hyperintense lesions in the basal ganglia
- E. Ischemic stroke with restricted diffusion in the left hemisphere

Answer: B

Verifier response: **Image-Text Consistency** NOT PASS: Image does not show bony lysis.

"Image-Text Consistency" means "Image and text align".



Mark only one oval.

- ☐ Yes, I agree with verifier's idea.
- ☐ Other: _____

This content is neither created nor endorsed by Google.

Google Forms