**Original Manuscript ID:** Access-2024-41310

**Original Article Title: “**Visibility Aware In-Hand Object Pose Tracking in Videos with Transformers”

**To:** IEEE Access Editor

**Re:** Response to reviewers

Dear Editor,

Thank you for allowing a resubmission of our manuscript, with an opportunity to address the reviewers’ comments.

We are uploading (a) our point-by-point response to the comments (below) (response to reviewers, under “Author’s Response Files*”*), (b) an updated manuscript with yellow highlighting indicating changes (as “Highlighted PDF*”*), and (c) a clean updated manuscript without highlights (“Main Manuscript”*).*

Best regards,

Dinh-Cuong Hoang, et al.

**Reviewer#1, Concern # 1:** The paper are well organized and ready for publication.

**Author response:**

**Author action:**

**Reviewer#2, Concern # 1:** Comparisons with State-of-the-Art Methods

**Author response:**

**Author action:**

**Reviewer#2, Concern # 2:** While the manuscript demonstrates improvements over existing methods, the comparison lacks depth. Consider expanding your discussion to include additional baselines or related work.

**Author response:** Thank

**Author action:**  We

**Reviewer#2, Concern # 3:** Some equations and the overall algorithm are not adequately described. Ensure that each mathematical element is defined and that the algorithm is presented with enough detail for reproducibility.

**Author response:** Thank

**Author action:**  We

**Reviewer#2, Concern # 4:** The methodology omits several key details, such as specific hyperparameters used, dataset preprocessing steps, and training settings. Providing this information would enhance the clarity and reproducibility of your work.

**Author response:** Thank

**Author action:**  We

**Reviewer#2, Concern # 5:** Some figures and plots lack sufficient annotations or context. Enhance these visual aids to ensure they are self-explanatory and clearly support the narrative.

**Author response:** Thank

**Author action:**  We

**Reviewer#2, Concern # 6:** Discuss the computational demands of your approach, particularly with regard to the transformer architecture, and propose strategies for real-time implementation.

**Author response:** Thank

**Author action:**  We

**Reviewer#2, Concern # 7:** Provide a clearer distinction between your visibility-aware module and similar mechanisms in the literature to highlight the unique contributions of your work.

**Author response:** Thank

***Author action:***  *We*

**Reviewer#2, Concern # 8:** While the references are generally relevant and up-to-date, consider including additional recent works that directly address visibility-aware tracking or transformer-based object pose estimation to strengthen the contextual foundation.

- While the manuscript is written in readable English, there are minor grammatical errors and phrasing inconsistencies. A thorough language review is recommended.

- A section discussing the limitations of your approach, such as scalability to larger datasets or performance in real-time scenarios, would provide a more balanced and transparent evaluation of your work.

**Author response:** Thank

***Author action:***  *We*