https://doi.org/10.59476/ilpmt2026.v2i1.792

COMPARATIVE ANALYSIS OF EYE-TRACKING AND AI TOOLS FOR PREDICTING VISUAL ATTENTION IN VIRTUAL HUMAN PERCEPTION

Gala Golubović, Sandra Dedijer, Saša Petrović, Sanja Cvetojević, Neda Milić Keresteš, Teodora Gvoka
University of Novi Sad, Serbia

Abstract

Visual attention plays an important role in how users perceive and engage with digital content, particularly in areas such as human–computer interaction and interface design. Eye-tracking is commonly used to study these processes, as it provides detailed and reliable data on visual attention. However, it requires specialized equipment and controlled experimental conditions, which may limit its practical application.

In recent years, artificial intelligence (AI) tools have been developed to predict visual attention based on previously collected eye-tracking data. While these tools are fast and easy to use, it remains unclear how closely their predictions reflect actual human behavior, particularly when observing virtual humans.

This paper presents a comparative study of eye-tracking results and AI-based attention predictions. The study involved 300 participants, whose visual attention was recorded while viewing a set of virtual human stimuli. The same stimuli were then analyzed using the AI tool across several available modes. The comparison was performed using SSIM, MSE, and Pearson correlation, as well as through the analysis of predefined areas of interest.

The results show a strong overall agreement between the two approaches, especially in identifying key facial regions such as the eyes and lips. At the same time, noticeable differences were observed in less prominent areas of the face, as well as during shorter observation periods. AI-generated maps also tend to appear more diffuse compared to eye-tracking data.

These findings suggest that AI tools can provide a useful approximation of visual attention patterns, particularly in the early stages of design. However, they do not fully capture the precision of eye-tracking and should therefore be used as a complementary method rather than a replacement.

Keywords: *eye-tracking, AI predictions, visual attention, virtual human*

Introduction

Understanding visual attention is important in areas such as user interface design, marketing, and human–computer interaction. In the past, researchers have used eye-tracking technology to study it. This method enables precise tracking of eye movements, but requires substantial resources and participant involvement (Duchowski, 2017; Holmqvist, 2011).

As artificial intelligence has advanced, researchers have created attention prediction tools using machine learning and deep learning models trained on eye-tracking data which offer fast and cost-efficient analysis (Kümmerer, Theis & Bethge, 2014; Cornia et al., 2018). However, it is still unclear how reliable they are compared to real eye-tracking data.

This question is especially important when studying virtual humans, since visual attention affects how people perceive emotions and interact with them (Iskra & Tomc, 2016; Calvo & Nummenmaa, 2008).

This study compares eye-tracking results with AI-based attention predictions when people observe virtual humans. The aim is to assess the degree of correspondence between the two methods.

Research Question: Can AI tools reliably reproduce patterns of visual attention obtained through eye-tracking methods?

Hypothesis: There is a significant correspondence between eye-tracking data and AI-based results.

Literature Review

Visual attention is a cognitive process involving the selective distribution of mental resources to relevant stimuli while simultaneously filtering out irrelevant information. In the context of face perception, research shows that the eye and lip regions are the most critical for gathering information about emotions and intentions (Iskra & Tomc, 2016; Calvo & Nummenmaa, 2008; Golubović et al., 2026).

Eye-tracking technology enables the direct measurement of visual attention through the analysis of eye movements, including fixations and saccades, providing objective and precise data on observer behavior (Duchowski, 2017; Holmqvist, 2011). This method is widely applied in fields such as psychology, marketing, and user experience design. In contrast, AI-based attention prediction models rely on machine learning and deep learning algorithms trained with large eye-tracking datasets. These models replicate visual scanning patterns and generate saliency maps indicating probable areas of focus (Kümmerer, Theis & Bethge, 2014; Cornia et al., 2018).

Modern AI tools are accurate and can quickly analyze visual attention without needing participants. However, research shows that these tools cannot fully replace eye-tracking technology and are mostly used to support it (Lavdas et al., 2021; Lengyel et al., 2021; Šola et al., 2024).

When creating virtual humans, it is important to understand how people pay attention to them. This helps designers make avatars look more realistic and improves the user experience in digital settings.

Methodology

The study is based on a comparative analysis of results obtained through eye-tracking and an AI-based attention prediction tool. The aim was to examine the degree of correspondence between empirically recorded gaze patterns and predictions generated by artificial intelligence models.

Participants and Stimuli

A total of 300 participants took part in the eye-tracking experiment. The stimuli consisted of 12 virtual humans (6 female and 6 male), in which only facial features were varied, while the background and clothing were standardized across all stimuli (Fig. 1).

Participants observed the presented characters and evaluated them according to three parameters: attractiveness, naturalness, and trustworthiness.

Fig. 1. Stimuli

Eye-Tracking Procedure

Data were collected using the GazePoint GP3 eye-tracker. During the experiment, participants' eye movements were recorded, and opacity maps were generated to provide a cumulative visualization of attention distribution.

To analyze the data in more detail, each character's face was split into areas of interest (AOIs). This allowed for a closer look at how attention was spread across different parts of the face (Fig. 2).

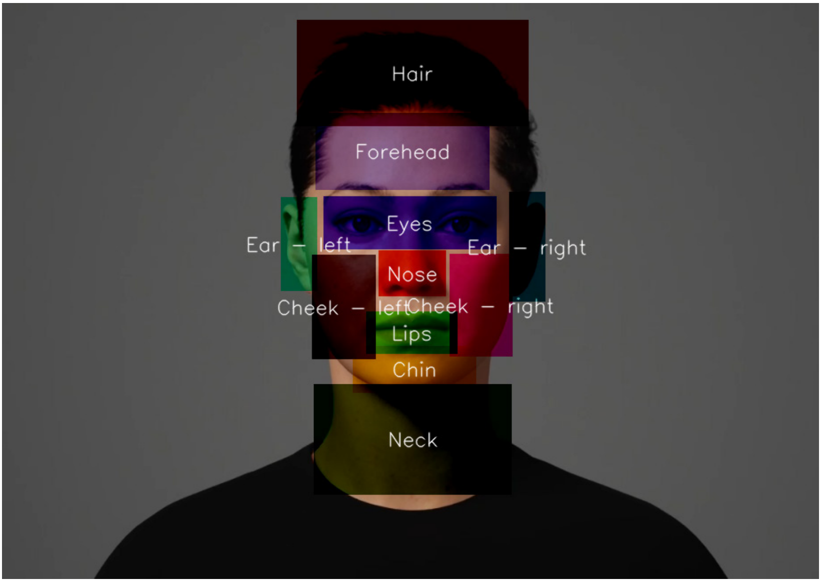


Fig. 2. Areas of Interest generated by GazePoint Analysis Software

For each character, opacity maps were made for three observation times: 3, 4, and 5 seconds. These maps show combined data from all participants (Fig. 3).

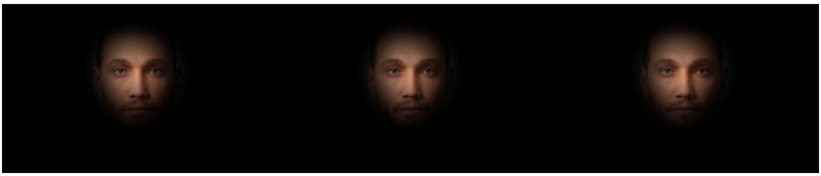


Fig. 3. Opacity maps for times of 3s, 4s, and 5s

AI-Based Attention Analysis

Used AI tool predicted visual attention by generating maps from deep learning models trained on eye-tracking data. The analysis was conducted using the same stimuli as in the eye-tracking experiment, and to capture a broader range of predictions, all available analysis modes offered by the software (*desktop, mobile, marketing material, packaging, poster*, and *shelves*) were used.

Focus maps, which most closely resemble eye-tracking opacity maps, served as the primary basis for comparison (Fig. 4).

Fig. 4. Focus maps in the Attention Insight software

Areas of interest (AOIs) were defined to match those in the eye-tracking analysis, allowing direct comparison of results (Fig. 5).

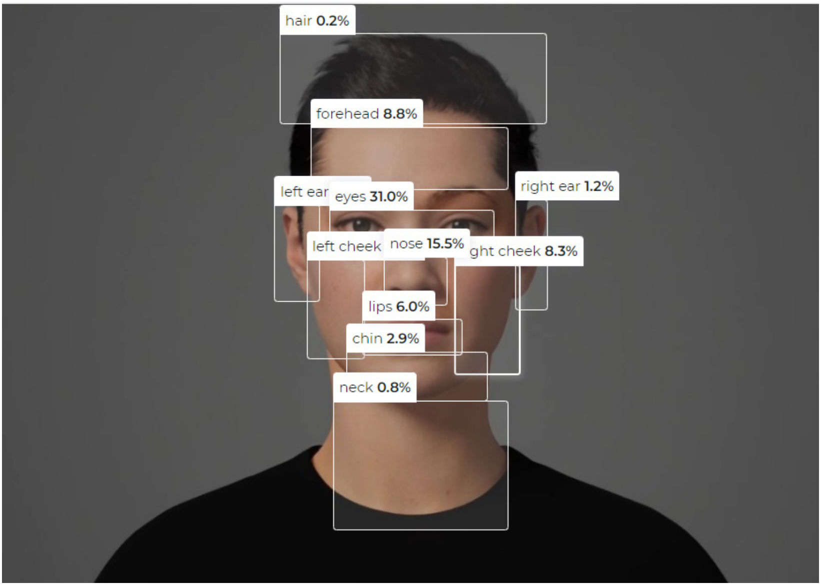


Fig. 5. Areas of Interest generated by AI software

Comparison Metrics

The following metrics were used for the quantitative assessment of correspondence between eye-tracking and AI-generated results:

- SSIM (Structural Similarity Index) – a measure of similarity between two images that accounts for structure, contrast, and luminance. Higher values indicate greater similarity.
- MSE (Mean Squared Error) – represents the average squared difference between pixel values of two images. Lower values indicate smaller error.
- Pearson correlation coefficient – measures the linear relationship between pixel values of two images. Values closer to 1 indicate a stronger positive correlation.

All analyses were conducted in MATLAB software, where comparisons were performed between opacity maps and AI-generated maps across all observation time intervals.

Types of Maps and Representation Selection

Two types of visual attention representations were considered: heat maps and opacity/focus maps. Heat maps provide an intuitive visualization of attention intensity through color gradients (Fig. 6), though their interpretation may differ across software platforms.

Therefore, opacity/focus maps were chosen for comparative analysis, ensuring clearer delineation of areas of interest and more consistent comparison across datasets.

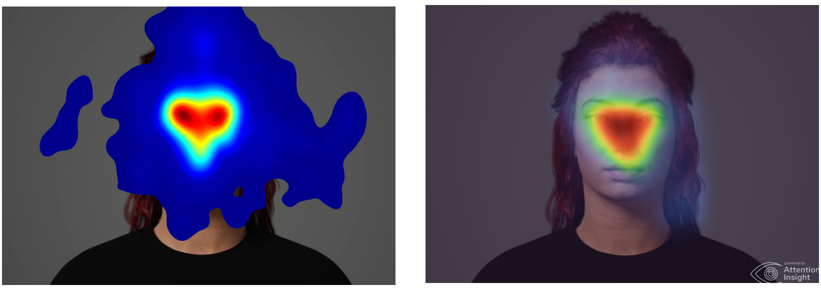


Fig. 6. Heat maps generated by GazePoint Analysis and Attention Insight

Results

The comparison between generated opacity maps and focus maps was conducted in MATLAB software. The analysis involved pairwise comparisons of results from each analysis mode in the AI tool with corresponding results from the GazePoint Analysis software. In total, 180 individual comparison pairs were generated (15 pairs for each of the 12 characters). The analysis included all stimuli and observation intervals (3s, 4s, and 5s) and used quantitative metrics such as SSIM, MSE, and the Pearson correlation coefficient.

Visual analysis revealed a substantial overlap in attention areas between eye-tracking and AI-generated results. AI models successfully identified key areas of interest; however, in some cases, deviations were observed in boundary precision and attention intensity, particularly in peripheral facial regions.

The quantitative analysis indicates a high degree of similarity between maps generated by the two methods:

- SSIM values indicate moderate to high structural similarity, confirming that AI models successfully reproduce global attention patterns.
- MSE values show relatively low error between maps, additionally supporting the quantitative closeness of the results.
- The Pearson correlation coefficient demonstrates a positive and statistically significant relationship between pixel values, indicating consistency in attention distribution across methods.

It was also observed that the degree of correspondence increases with observation time (from 3s to 5s), suggesting that AI models better approximate stabilized attention patterns than the initial phases of visual scanning.

Regarding the analysis modes within AI tool, the *shelves* mode demonstrated the highest level of correspondence, specifically when evaluated with the Structural Similarity Index (SSIM). This suggests that the *shelves* mode produces results most closely aligned with human visual perception according to SSIM. In comparison, other modes such as *mobile* and *packaging* did not achieve as high SSIM values but ranked differently across other metrics.

When evaluated using Mean Squared Error (MSE), the best performance was observed in the *mobile* mode, indicating it had the highest level of objective (mathematical) correspondence. Despite this, *mobile* mode may not appear most similar to human observers. Notably, the *shelves* mode ranked second in MSE-based similarity, demonstrating that it outperformed other modes except mobile mode.

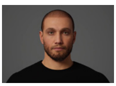
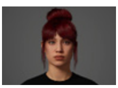
According to the Pearson correlation coefficient, the highest correlation was observed in the *packaging* mode, suggesting the strongest relationship in pixel value trends for this mode. The lowest correlation was associated with the *marketing material* mode. This suggests that the *packaging* mode shows the strongest correspondence for this metric.

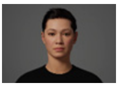
Observation Time

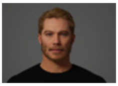
Since both software tools support defining areas of interest (AOIs) and recording the percentage of attention directed to each region, an additional comparison was made using these metrics. As the strongest alignment between eye-tracking results and AI predictions emerged at the 5-second interval, this analysis focused solely on data from that period.

The comparison of these values showed that predictions generated with the *packaging* mode most closely matched recorded gaze behavior. The *shelves* mode was the next most similar (Table 1).

Table 1. Proportion of total observation time allocated to the area of interest

		GazePoint Analysis	**Attention Insight**				
Virtual Human	AOI	Ave Time (%)	Desktop (%)	Marketing material (%)	Mobile (%)	Packaging (%)	Shelves (%)
	Eyes	41.7	21.9	21.6	21.8	21.3	**26.1**
	Nose	11.9	15.5	25.6	16.4	**9.1**	18.9
	Lips	7.4	8.3	10.1	12.5	6.3	**6.7**
	Hair	4.8	0.0	0.0	0.1	**2.2**	0.0
	Neck	3.2	4.0	**3.1**	2.5	2.7	1.7
	Left ear	1.2	1.4	0.0	0.7	4.3	**1.1**
	Right ear	1.3	**1.8**	0.1	2.0	**1.8**	0.6
	Chin	3.2	10.8	8.3	9.5	6.6	**5.9**
	Left cheek	4.4	6.5	**3.5**	6.5	6.9	9.0
	Right cheek	2.6	6.9	**5.9**	9.0	6.6	7.9
	Forehead	14.2	7.9	5.6	9.7	**11.7**	9.0
	Eyes	37.3	22.4	**35.4**	22.1	18.5	21.9
	Nose	10.9	16.0	33.1	17.6	**9.4**	16.4
	Lips	5.1	10.5	8.8	10.1	**6.8**	7.3
	Hair	7.2	4.9	0.7	2.9	2.0	**9.1**
	Neck	3.4	**2.9**	0.7	1.0	0.6	0.4
	Left ear	2.3	0.1	0.1	0.9	**2.7**	0.6
	Right ear	1.2	1.0	0.4	2.4	2.5	1.2
	Chin	3.1	5.5	**3.0**	5.0	4.4	3.6
	Left cheek	2.5	4.7	**2.3**	6.2	8.1	5.7
	Right cheek	2.4	5.3	**4.4**	5.2	7.1	6.9
	Forehead	16.0	8.5	9.8	**12.5**	9.6	10.6

	Eyes	37.6	21.5	20.3	22.5	12.8	**22.8**
	Nose	8.0	13.4	28.7	13.8	**10.4**	16.0
	Lips	5.2	8.4	9.1	8.6	7.7	**6.6**
	Hair	8.2	0.3	0.1	1.0	**1.3**	0.5
	Neck	3.3	4.3	0.5	2.4	6.1	**2.6**
	Left ear	1.2	1.5	0.0	**1.4**	3.4	2.0
	Right ear	1.7	2.4	0.6	**2.4**	2.5	3.4
	Chin	2.8	7.5	**4.2**	8.8	9.7	4.3
	Left cheek	3.3	5.9	**3.5**	6.6	5.8	6.7
	Right cheek	3.3	7.2	**6.8**	8.5	6.8	10.5
	Forehead	20.1	7.1	3.9	9.1	7.5	**11.2**
	Eyes	36	21.8	**27.9**	19.3	22.1	27.5
	Nose	10.3	14.3	19.0	17.3	**9.4**	20.7
	Lips	4.6	7.3	9.1	9.7	**4.0**	9.0
	Hair	4.7	0.0	0.0	0.0	**3.4**	0.3
	Neck	3.3	1.0	0.4	**2.8**	0.7	0.4
	Left ear	2.9	0.8	0.0	0.5	**4.1**	1.3
	Right ear	0.9	2.3	0.2	**1.1**	2.3	**1.1**
	Chin	4.6	5.8	2.6	**5.5**	3.6	3.1
	Left cheek	3.8	6.6	1.3	**4.3**	8.0	6.2
	Right cheek	5.2	7.7	3.8	7.1	6.2	**5.3**
	Forehead	19.3	12.3	8.2	7.0	**16.1**	11.2
	Eyes	40.2	22.9	27.3	23.5	21.2	**31.0**
	Nose	8.0	15.6	24.5	14.5	**11.8**	15.5
	Lips	2.6	6.3	9.8	7.1	**5.2**	6.0
	Hair	5.8	0.2	0.0	1.0	**2.1**	0.2
	Neck	2.5	7.8	1.1	**2.8**	**2.2**	0.8
	Left ear	1.2	**1.2**	0.0	0.9	4.5	2.0
	Right ear	1.0	2.1	0.4	2.4	2.2	**1.2**
	Chin	2.0	8.1	3.5	8.1	5.0	**2.9**
	Left cheek	1.8	7.4	4.4	9.9	6.5	**4.3**
	Right cheek	2.4	5.6	**5.4**	9.3	6.8	8.3
	Forehead	16.5	8.0	7.2	**11.2**	**11.2**	8.8
	Eyes	37.9	20.1	**27.4**	19.0	21.1	22.9
	Nose	8.5	15.5	31.5	18.1	**12.5**	19.3
	Lips	4.2	9.7	12.4	10.7	**6.7**	8.3
	Hair	14.7	5.0	1.5	5.8	**15.2**	7.0
	Neck	2.8	5.1	1.4	1.8	5.2	**2.8**
	Chin	2.2	7.4	**3.6**	5.5	4.8	4.1
	Left cheek	2.3	7.8	**3.4**	4.6	8.5	7.7
	Right cheek	2.7	7.0	**5.3**	5.9	9.5	8.7
	Forehead	22.2	8.9	4.3	**10.0**	8.8	6.9

	Eyes	24.7	19.9	28.8	19.9	17.7	**23.7**
	Nose	9.0	17.7	20.6	16.4	**9.8**	14.0
	Lips	4.5	11.5	8.5	8.7	**6.7**	9.6
	Hair	7.4	0.3	0.1	0.8	**3.6**	2.2
	Neck	2.5	4.5	**2.1**	2.0	0.4	0.5
	Left ear	1.7	0.2	0.0	0.3	2.9	**0.6**
	Right ear	1.4	1.6	0.4	2.2	2.1	**1.2**
	Chin	2.0	10.4	6.5	9.5	**5.5**	7.8
	Left cheek	2.5	4.6	**1.8**	4.4	7.1	4.4
	Right cheek	2.7	6.7	**4.5**	5.6	5.6	7.0
	Forehead	26.6	12.2	8.8	**16.8**	13.8	11.0
	Eyes	31.0	20.6	**28.1**	21.3	17.8	24.5
	Nose	8.9	11.0	24.2	16.0	**10.7**	14.1
	Lips	4.3	8.4	7.0	9.7	4.9	**4.0**
	Hair	7.3	0.2	0.0	0.8	**4.6**	1.5
	Neck	2.7	**3.2**	0.7	**2.2**	1.2	0.0
	Left ear	2.7	1.0	0.1	1.3	3.9	**1.8**
	Right ear	1.4	2.5	0.8	**1.9**	3.2	2.1
	Chin	2.3	5.6	4.3	5.8	5.2	**2.7**
	Left cheek	3.5	7.4	**2.8**	6.7	8.6	7.1
	Right cheek	3.9	7.3	**6.4**	8.9	7.3	8.7
	Forehead	22.9	11.2	7.5	**15.2**	11.5	10.6
	Eyes	36.8	22.8	**24.8**	21.5	15.4	24.2
	Nose	10.1	15.1	22.8	18.1	**13.2**	16.8
	Lips	8.5	10.3	10.3	11.4	7.8	**8.6**
	Hair	4.0	0.2	0.1	0.9	**1.9**	0.2
	Neck	2.5	**2.7**	1.1	**2.7**	0.0	1.3
	Left ear	3.9	0.2	0.0	0.4	**3.1**	1.1
	Right ear	3.5	1.4	0.1	1.2	**2.0**	0.9
	Chin	4.1	9.3	**6.1**	8.4	7.7	6.8
	Left cheek	3.6	4.5	2.3	**2.9**	10.3	5.2
	Right cheek	4.1	5.7	3.4	**4.5**	6.5	7.5
	Forehead	15.5	10.3	6.0	**14.6**	12.0	12.3
	Eyes	28.4	20.3	25.1	23.1	22.7	**25.2**
	Nose	15.6	18.6	25.2	**16.3**	22.4	18.5
	Lips	7.5	2.3	13.1	10.4	11.0	**8.8**
	Hair	14.5	3.7	1.1	**7.7**	3.5	6.9
	Neck	3.7	**3.2**	0.9	0.7	1.8	1.8
	Chin	2.9	6.3	3.9	5.0	4.9	**2.8**
	Left cheek	2.3	3.8	**2.2**	6.3	5.3	5.5
	Right cheek	4.9	**5.3**	**5.3**	5.4	8.3	6.7
	Forehead	21.2	9.5	5.9	**10.4**	7.8	5.9

	Eyes	28.4	22.2	30.3	19.0	20.8	**27.3**
	Nose	15.6	**16.1**	26.5	19.2	10.4	13.2
	Lips	7.5	**8.7**	8.8	9.5	4.8	5.1
	Hair	4.9	7.1	2.0	**5.0**	15.3	12.7
	Neck	21.2	**3.6**	0.7	2.6	1.2	1.3
	Chin	3.7	6.4	**3.3**	4.2	2.3	2.0
	Left cheek	14.5	4.3	2.1	6.4	**7.8**	4.3
	Right cheek	2.3	**5.3**	6.4	5.7	8.1	7.3
	Forehead	2.9	8.4	**6.5**	11.9	11.0	9.8
	Eyes	27.3	21.7	**32.5**	21.4	16.4	21.4
	Nose	15.1	**17.0**	28.3	19.4	10.1	19.6
	Lips	5.4	10.0	**7.2**	9.5	7.6	8.0
	Hair	6.1	**9.1**	0.0	0.6	2.5	0.0
	Neck	4.0	2.7	1.9	1.1	**3.5**	1.6
	Left ear	2.2	**0.7**	0.0	0.6	4.1	**0.7**
	Right ear	2.4	1.6	0.5	1.4	**2.7**	1.5
	Chin	3.8	6.6	3.2	6.5	8.4	**3.8**
	Left cheek	6.1	**6.1**	2.5	3.9	7.7	4.8
	Right cheek	6.1	6.8	4.2	5.3	**6.2**	7.0
	Forehead	13.7	2.7	5.9	**12.3**	12.2	10.1

Observation Time in Relation to Character Gender

When comparing the average observation time for each area of interest (AOI) by character gender, slight differences were observed. The most notable difference was observed in the viewing of hair, where the average observation time for female characters' hair was approximately twice as long as that for male characters (Table 2).

Furthermore, in the case of male characters, more attention was allocated to the beard and cheek regions. Apart from the eye region – which consistently attracts the highest level of attention – these findings suggest that, compared to other facial areas, greater attention is directed toward the lower part of the face when observing male characters.

In contrast, for female characters, attention is more frequently focused on the upper part of the face.

When comparing the average observation times from GazePoint Analysis with the averaged predictions from AI tool, it was found that for male characters, the predictions show the highest correspondence in the *packaging* mode. For female characters, the *marketing material* mode shows a slight advantage in correspondence.

Table 2. Observation time based on the Virtual Humans' gender

AOI	Male Virtual Human (s)	Female Virtual Human (s)
Eyes	32.9	35.4
Nose	10.9	10.2
Lips	5.8	5.0
Hair	5.7	11.3
Neck	3.0	3.3
Left ear	2.4	1.6
Right ear	1.8	1.3
Chin	3.3	2.5
Left cheek	4.0	2.6
Right cheek	4.1	3.6
Forehead	18.7	19.2

Key Findings

Based on the conducted analyses, the following key findings can be highlighted:

- AI models effectively detect the primary regions of visual attention on virtual characters.
- A strong positive correlation exists between AI-predicted attention maps and results from traditional eye-tracking, indicating substantial agreement in measured visual attention patterns between the two methods.
- The largest differences appear in the outer areas of the stimuli.
- AI models become more accurate when observation times are longer.
- AI models usually show a less precise, more blurred pattern of attention than eye-tracking data.

These results show that AI tools can reliably capture overall patterns of visual attention, though they have some limits in precision.

Discussion

This study aimed to examine whether AI-based attention prediction tool can reliably replicate the visual attention behaviors found with eye-tracking when people observe virtual humans.

The results indicate a strong correspondence between the two methods, especially in spotting main areas of attention. This is consistent with previous research that points out the importance of these regions in face perception (Iskra & Tomc, 2016; Calvo & Nummenmaa, 2008). It confirms that AI models can capture key aspects of visual attention.

However, the differences found, especially in the outer facial areas and in how precisely attention is mapped, show the limits of AI-based methods. Eye-tracking records real user behavior as it happens, while AI models make predictions based on patterns from data. Because of this, AI models give a simpler and more spread-out picture of attention.

One key finding is that the match between the two methods improves with longer observation times. This suggests that AI models perform better at predicting steady phases of visual attention, but have trouble with the quick, early eye movements that often happen without conscious thought. This may be because early visual attention is more affected by personal and situational factors, which are hard for AI models to generalize.

In practice, these results indicate that AI tools can be a useful and effective alternative to eye-tracking in the early stages of design, when quick and affordable checks of user attention are needed. They are especially helpful in areas like UI/UX design, marketing, and virtual human development.

However, for studies that need a very precise and detailed analysis of behavior, eye-tracking is still essential. In this way, AI tools should be seen as a complement to, not a replacement for, eye-tracking, which corresponds to the results of earlier studies (Lavdas et al., 2021; Lengyel et al., 2021; Šola et al., 2024).

The results partly confirm the hypothesis: there is a strong match between AI and eye-tracking results, but there are clear limits in how precise and interpretable the AI results are.

Conclusion

The findings show a high level of agreement between the two methods, suggesting that AI tools can closely approximate overall patterns of visual attention. However, some limitations were observed, particularly in the precision of attention mapping, the analysis of peripheral areas of the stimuli, and the prediction of early observation phases.

For this reason, AI tools cannot fully replace eye-tracking, but they can be considered a useful complementary method for studying visual attention.

These results also point to the practical value of AI tools in the early stages of design, where quick and cost-effective evaluation of user attention is needed, especially in UI/UX design, marketing, and virtual character development.

Directions for future research include:

- The use of multiple AI tools and their combinations
- Analysis of dynamic stimuli (video and animation)
- Inclusion of additional metrics and biometric data
- Investigation of the influence of individual user characteristics on attention patterns

Further improvements in AI models, particularly through integration with real-world data, could help reduce existing differences and support their broader application.

Acknowledgement

This research has been supported by the Ministry of Science, Technological Development and Innovation (Contract No. 451-03-34/2026-03/200156) and the Faculty of Technical Sciences, University of Novi Sad through project "Scientific and Artistic Research Work of Researchers in Teaching and Associate Positions at the Faculty of Technical Sciences, University of Novi Sad 2026" (No. 01-3609/1).

List of references

1. Calvo, M. G., & Nummenmaa, L. (2008). Detection of emotional faces: Salient physical features guide effective visual search. Journal of Experimental Psychology: General, 137(3), 471–494. https://doi.org/10.1037/a0012771
2. Cornia, M., Baraldi, L., Serra, G., & Cucchiara, R. (2018). Predicting human eye fixations via an LSTM-based saliency attentive model. IEEE Transactions on Image Processing, 27(10), 5142–5154. https://doi.org/10.1109/TIP.2018.2851672
3. Duchowski, A. T. (2017). Eye tracking methodology: Theory and practice. Springer.
4. Golubović, G., Dedijer, S., Keresteš, N. M., Pál, M., Kerac, J., & Đurđević, S. (2026). On the influence of virtual characters' facial features on young adult user perception: A quantitative study. International Journal of Humanoid Robotics. Advance online publication. https://doi.org/10.1142/S0219843626500015
5. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). Eye tracking: A comprehensive guide to methods and measures. Oxford University Press.
6. Iskra, A., & Tomc, H. G. (2016). Eye-tracking analysis of face observing and face recognition. Journal of Graphic Engineering and Design, 7(1), 5–11. https://doi.org/10.24867/JGED-2016-1-005

7. Kümmerer, M., Theis, L., & Bethge, M. (2014). Deep gaze I: Boosting saliency prediction with feature maps trained on ImageNet. arXiv preprint. https://doi.org/10.48550/arXiv.1411.1045
8. Lavdas, A. A., Salingaros, N. A., & Sussman, A. (2021). Visual attention software: A new tool for understanding the "subliminal" experience of the built environment. Applied Sciences, 11(13), 6197. https://doi.org/10.3390/app11136197
9. Lengyel, G., Carlberg, K., Samad, M., & Jonker, T. R. (2021). Predicting visual attention using the hidden structure in eye-gaze dynamics. In CHI 2021 Eye Movements as an Interface to Cognitive State (EMICS) Workshop Proceedings. ACM.