注視点に基づいた個人識別への顕著性マップ手法導入の検討

長沢 潤一[1,2,a)]　礒本 俊弥[2,b)]　廣江 葵[4,3,c)]　中田 裕一[1,3,d)]　長松 隆[3,e)]

概要：情報セキュリティの重要性が高まる中，従来の認証方法には初回認証後のセキュリティ維持等の課題が存在する．デバイスを頻繁にロックさせることによりセキュリティを強化することはできるが，一方で定期的な再認証要求はユーザビリティを著しく低下させる．このジレンマを解決するため，継続的かつ暗黙的な認証方法の開発が求められる．本研究では，この課題に対応するため，継続的かつ暗黙的な認証方法の開発を目的とする．予備調査として，注視点に基づく個人識別の可能性を探求するため，公開データセットである UEyes から 10 名の実験参加者を抽出し，15 枚の画像に対する視線ヒートマップを比較分析した．実験参加者間の視線パターンを分析した結果，個人間の眼球運動に同程度の差異が存在することが示された．この結果は，眼球運動特性を利用した個人識別が可能であることを示唆している．今後の研究では，個人の視線に特化した顕著性マップの開発など，より高度な個人識別手法の開発を目指す．本研究は，ユーザのセキュリティとユーザビリティを両立させるために，眼球運動に着目した新たな認証方法の基礎となる可能性を示す．

1. はじめに

情報の機密性を保護するためには，情報端末へのアクセス制限を行う必要がある．日常生活で採用されている認証方法は数多くあるが，パスワード認証が依然として広く使われている [1]．パスワード認証の利点の 1 つは，任意のパターンを決定でき変更できることであるが，他人に知られると悪用される可能性がある．パスワード認証を使用しない認証方法として，顔認証や指紋認証に代表されるバイオメトリクス認証がある．この認証方法は，パスワードを使用した認証方法と異なり，ユーザが記憶する必要がなく，個人の身体的特徴を用いるため一度登録すれば容易に使用することができるという利点があるが，認証データが流出すると機密性の点から使用できなくなるという問題がある [2], [3]．

昨今，スマートフォンや PC，ヘッドマウントディスプレイ（HMD）などのデバイスでは，パスワード認証やバイオメトリクス認証が一般的に用いられている．しかし，最初の認証後に別ユーザがデバイスを使用する場合，デバイスはログインされたまま別ユーザがアクセス可能な状態になるため，セキュリティが損なわれる可能性がある．これは，ユーザがデバイスから離れる際にログアウトすることで解決できるが，都度ログアウト・ログインすることの煩わしさやログアウトを忘れるといったことも考えられる．さらなるセキュリティ担保のためには，頻繁な再認証の実施だけではなく，ユーザが変化したことを暗黙的に認識できる手法が望まれている．

この問題の解決策として，ユーザの連続的な行動を使用した認証が提案されている．Acar ら [4] は，キーストロークのダイナミクス，マウスの動き，タッチジェスチャなどのさまざまなモダリティを用いた，明示的な認証を必要しないユーザ認証を論じている．同様に，Ehatisham-ul-Haq ら [5] も，パッシブモバイルセンシングを使用した行動パターン認識に基づく，スマートフォンユーザ向けの認証システムを提案している．

HMD を使用する場合，認証には主にパスワードが使用されている．しかし Wang ら [6] は，目の動きからパスワードを含む入力テキストを推測できることを発見したため，HMD 内でパスワードを入力する認証方式にはセキュリティ上の問題があることが明らかとなった．この問題に対処するため，一定の入力パターンを使用することなく，視線データの特徴のみに基づいて個人を識別することが可能

1 関西学院大学
2 LINE ヤフー株式会社
3 神戸大学
4 大阪成蹊大学
a) jnagasawa@acm.org
b) r.t.isomoto@gmail.com
c) hiroe@ieee.com
d) ynakata@person.kobe-u.ac.jp
e) nagamatu@kobe-u.ac.jp

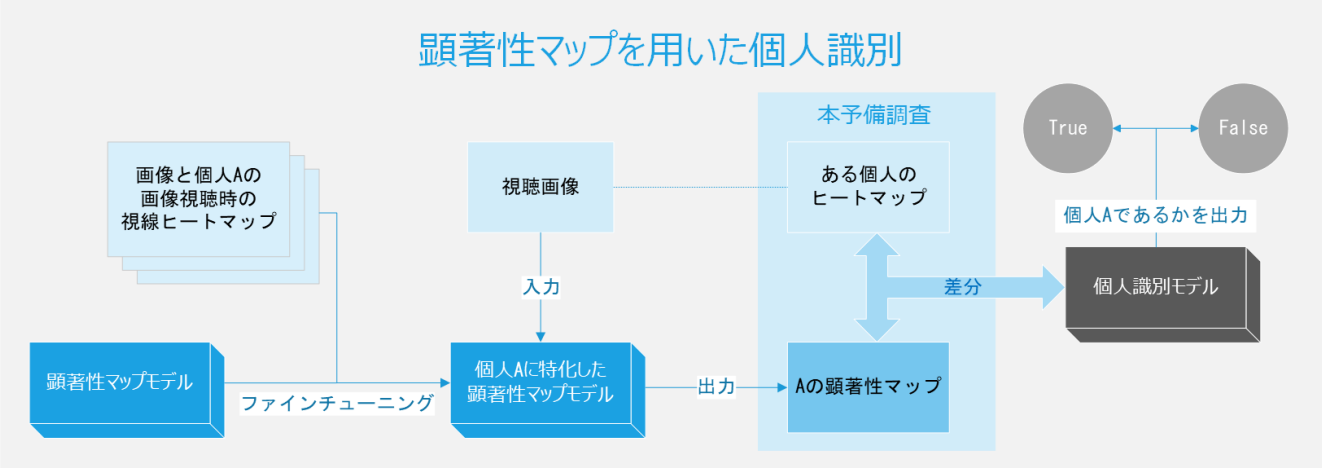

図 1 研究全体の概念図．青枠の部分は，本予備調査の範囲である．
Fig. 1 Schematic diagram of the overall research. The area in blue represents the scope of this preliminary study.

になれば，デバイスに内蔵されたセンサを使用して認証を行うことができる．さらに，ある画像を閲覧した際の視線パターンを推測・模倣することは困難であることが予想されるため，バイオメトリクス認証のような堅牢さも兼ね備えている．これにより，パスワード入力などの煩雑なプロセスが不要となり，ユーザの利便性を損なうことなく，リアルタイムかつ暗黙的な認証が可能となる．

本研究の目標は，**図 1** に示すような視線データの特徴のみから個人を暗黙的に識別するシステムの開発である．そのための予備調査として，本稿では図 1 の青枠内に示す箇所，すなわち個人間の視線パターンの差異の検出が可能かどうかを調査する．具体的には，ある画像を提示した際のユーザ毎の注視点の推移にどの程度の差異があるかを調べ，注視点を元に作成された顕著性マップ（画像の中で注意を引きやすい領域をヒートマップ形式で可視化したもの）を個人識別に使用することの実現可能性を検討する．この調査には，公開済みデータセットである UEyes[7]*1 を使用した．UEyes は，1,980 枚の画像を 62 名の参加者が閲覧した際の視線データを含む大規模な視線追跡ベースのデータセットである UEyes から共通の 15 枚の画像を見た 10 名の参加者を抽出し，画像閲覧時の参加者間のヒートマップを比較し，分析を行った．分析の結果，参加者間の視線データにはそれぞれ同程度の差異があることがわかった．

2. ヒートマップにおける個人差の調査

本研究では，UEyes データセットから共通の 15 枚の画像を閲覧した 10 名の参加者を抽出した．各参加者の画像閲覧時における視線座標データに対し，ガウシアンフィルタを適用して正規化を行い，視線ヒートマップを生成した．これらの視線ヒートマップを本研究の分析対象とした．顕著性マップは，MIT/Tuebingen Saliency Benchmark[8]–[10] の MIT300，CAT2000，COCO Freeview データセットで高いベンチマークスコアを獲得した DeepGaze IIE モデル[11] を用いて生成した．

2.1 結果

本研究において使用した 15 枚の画像のうち 1 枚の画像を使用した結果を例示する．**図 2** は閲覧した画像，画像に対応する顕著性マップ，および実験参加者の視線ヒートマップを示す．**図 3** は，図 2 に示された全実験参加者の視線ヒートマップの平均（mean）と標準偏差（SD）を表している．平均を示すヒートマップは濃い色であるほど参加者の注意が向けられていたことを示し，標準偏差を示すヒートマップは濃い色であるほど参加者毎に注意の向け方に違いがあったことを示す．図 2 の顕著性マップからは参加者の注意が主に画像左側の文字列と画像右側の女性の顔周辺に集まる可能性が高いことが分かる．そして，図 3 の視線ヒートマップの平均から，参加者が顕著性マップによって示された箇所周辺を実際に注視していることが多いとわかる．一方で各参加者毎の視線ヒートマップを観察すると，文字列に注意を向けていない（Participant 63）参加者および女性の顔に注意を向けていない（Participant 19）参加者がいたこともわかる．加えて，視線ヒートマップの標準偏差（図 3，右）からは，画像上部は参加者毎に注意の向け方にばらつきがあったことがわかる．同様に，参加者の注意が多く向けられた文字列および女性の顔に関しても，その向け方にはユーザ毎に違いがあることがわかる．

図 2 を閲覧した際に得られた各画素に対応した視線座標の相関を参加者間毎に計算した結果を**図 4** に示す．相関行列の各セルは，参加者同士の視線パターンの相関を示し，相関 1.0 で完全に一致，相関-1.0 で完全に不一致を意味す

*1 https://userinterfaces.aalto.fi/ueyeschi23/

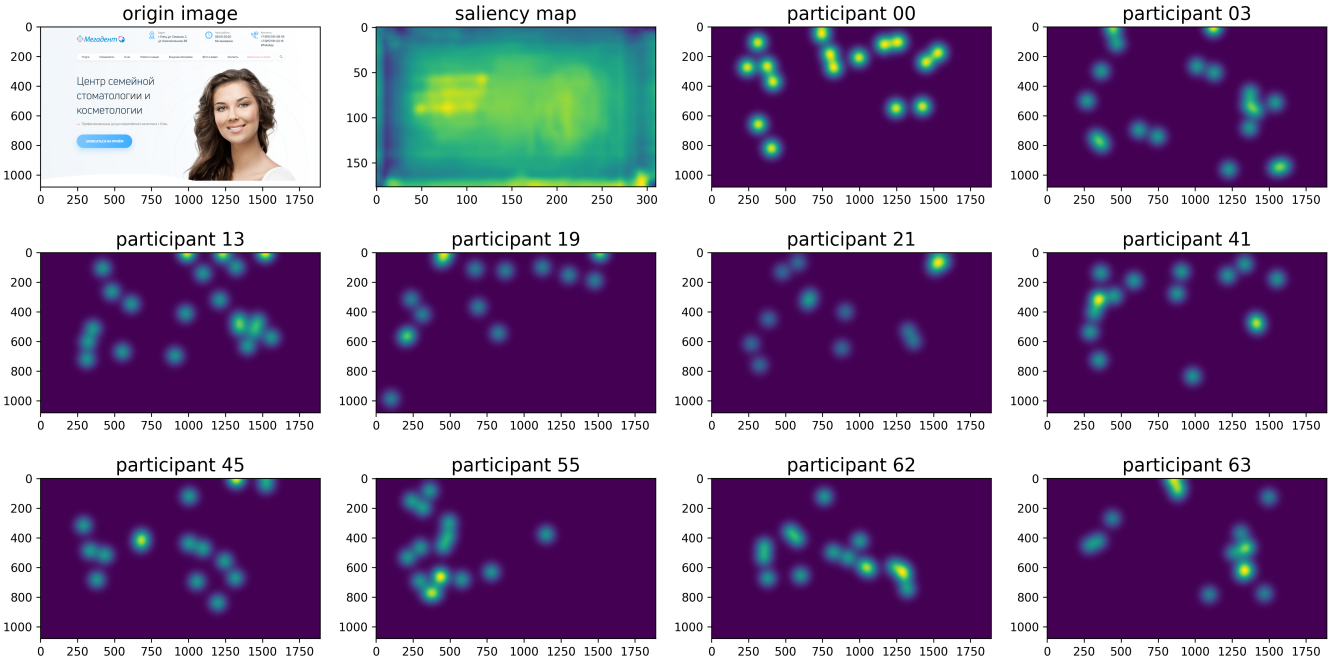

図 2 参加者が閲覧した画像（上段左），DeepGaze IIE で生成した顕著性マップ（上段左から 2 番目），および実験参加者の視線ヒートマップ（他の画像）の比較．

Fig. 2 Original image viewed by participants (top left), saliency map generated by DeepGaze IIE (second from left in the top row), and individual gaze heatmaps of participants while viewing the image (others).

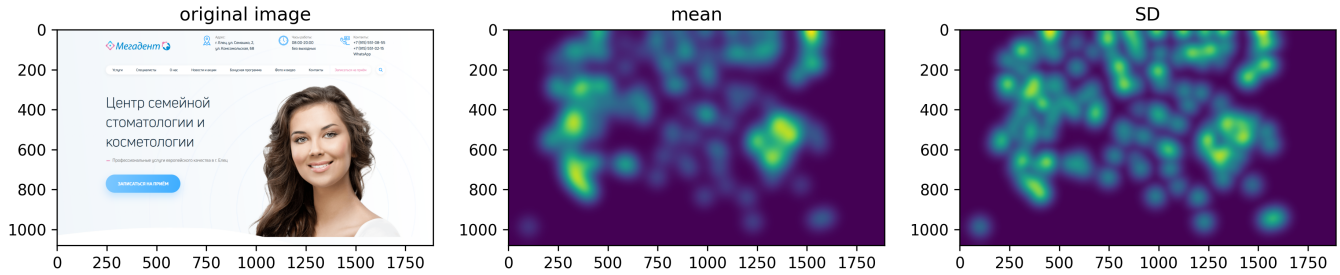

図 3 閲覧画像と参加者の視線データ比較．参加者が閲覧した画像（左），参加者全員の視線ヒートマップの平均（中央），参加者全員の視線ヒートマップの標準偏差（右）．

Fig. 3 Viewed Image and Comparisons of Participants' Gaze Data. Original image viewed by participants (left), average gaze heatmap of all participants (center), standard deviation of gaze heatmaps of all participants (right).

る．同じ画像を見た際の注視箇所はある程度一致することが予想されることから，概ねどのセルでも相関は 0.0 を上回ることが予想できる．多くの実験参加者ペアにおいて視線パターンにある程度の正の相関が見られるが，一部のペアでは負の相関も観察され，異なる視線パターンを持ったペアがいることがわかる．対照的に，15 枚の全画像の視線ヒートマップにおける実験参加者間の相関行列**図 5** では，全実験参加者ペア間で概ね均一な正の相関が見られ，視線パターンがある程度一致し，かつ同程度の差異が存在することがわかる．

2.2 考察

本研究では，ヒートマップ分析と相関行列を用いて，実験参加者の注意のパターンを調査した．

まずヒートマップの標準偏差の分析結果（図 3 右）からは，実験参加者間で視線にばらつきがあることがわかった．UEyes データセットの異なる画像に対するヒートマップの特徴も調査したところ，視線のばらつきの傾向が画像によって大きく異なることがわかった．具体的には，複雑な画像ほど標準偏差が高くなる傾向が観察されたが，これは画像の複雑さが増すにつれて，個人の興味や解釈の違いがより顕著になることが理由だと考えられる．これは，本研究が目的とする視線の差異を用いた個人識別の実現可能性を示唆する結果である．

個別の画像については，相関行列の分析から，実験参加者間の視線パターンが画像によって大きく異なることが明

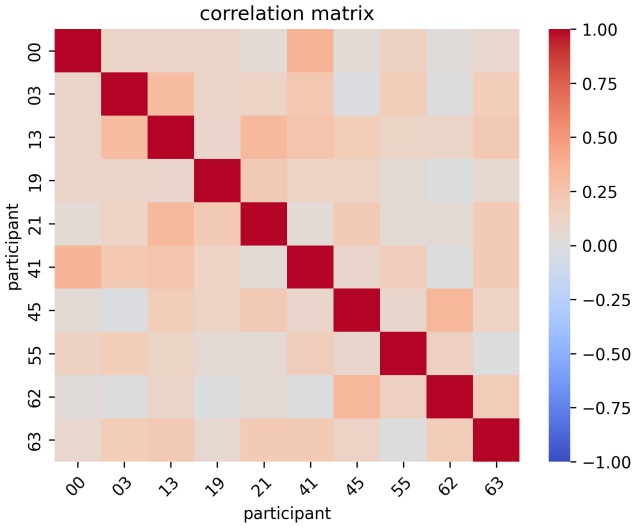

図 4 各画像に対する実験参加者ごとの視線ヒートマップの相関行列．参加者番号は UEyes データセットに沿った参加者番号を示す．各セルの色は相関の強さを表し，対角線上のセルは同一参加者間の完全な相関（1.0）を示す．

Fig. 4 Correlation matrices of gaze heatmaps between participants for each image. The color of each cell represents the strength of correlation, with diagonal cells showing perfect correlation (1.0) between identical participants.

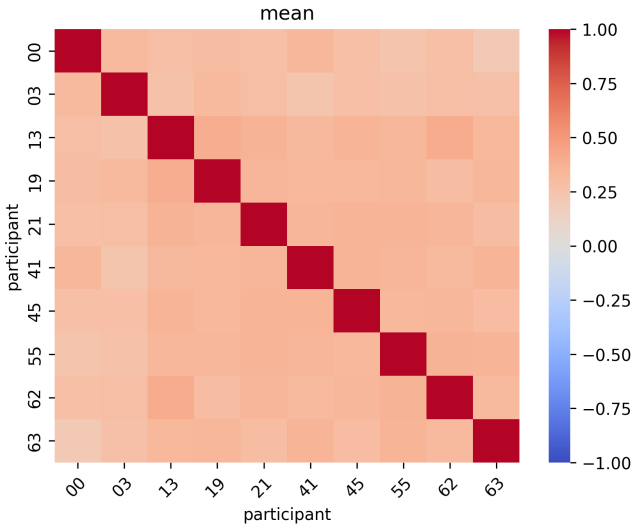

図 5 15 枚全ての画像における実験参加者ごとの視線ヒートマップの相関行列．参加者番号は UEyes データセットに沿った参加者番号を示す．各セルの色は相関の強さを表し，対角線上のセルは同一参加者間の完全な相関（1.0）を示す．この行列は参加者間の全体的な視線パターンの類似性を表している．

Fig. 5 Correlation matrix of gaze heatmaps for each participant across all images. The color of each cell represents the strength of correlation, with diagonal cells showing perfect correlation (1.0) between identical participants. This matrix represents the overall similarity of gaze patterns among participants.

らかになった．例えば，図 2 の画像（女性の顔と文字が含まれた画像）に対しては，実験参加者間の相関が低く，各実験参加者が画像の異なる部分に注目していたことがわかった．一方で，UEyes データセットの他の画像では参加者間の相関が高いものもあり，共通の注目領域があったことを推測できる．

図 4 および図 5 の相関行列において，正の相関（赤いセル）は参加者間の注視点の類似性を，負の相関（青いセル）は注視点の相違を示している．相関行列の各セルの値にばらつきが大きい場合，視線パターンが類似したペアとそうでないペアが混在することとなり，全てのユーザーを一貫して識別することが困難になる．そのため本研究が目的とする，注視点に基づいた個人識別を実現するには，参加者間の相関が強くない（参加者間のセルの色が濃すぎない）ことおよび，相関にばらつきがないことが理想的である．一枚の画像を閲覧した際の注視点のみに焦点を当てた場合は，図 4 が示すように各セルの相関にばらつきが見られる．一方で，15 枚の全ての画像の視線ヒートマップから作成した相関行列では，図 5 で示される通りどの参加者のペアでも相関は全セルにおいて平均 0.32（SD=0.04）と同程度である．この結果から，視線ヒートマップでは個人間の差異は画像によってばらつきがあるものの，画像数を増やすことでその差異のばらつき具合は小さくなることがわかった．これは今後実験を行う際に，様々なタスクおよび UI での実験実施の必要性を示している．

これらの考察から，画像を見ている際の視線には一定の差異が認められるため，本研究が目的とする注視点に基づいた個人識別の実現可能性はあると考える．一方で，特に画像自体に特徴が少ない物を見ている際には，個人差が発生しない可能性も考えられる．この点には十分に留意する必要があり，今後，ユーザが見ている画像が特徴的であるかを自動的に評価し，それによって個人識別のパラメータを適応的に調整することなども検討する．

3. まとめ

本研究では，ユーザの暗黙的な認証システムの開発に向けた予備実験として，眼球運動の個人差に着目し，ヒートマップを用いた分析を行った．UEyes データセットから抽出した 10 名の実験参加者の画像閲覧時のヒートマップを比較・分析することで，注視点の個人差が個人識別に有効である可能性を検証した．

研究の主な成果は以下の通りである：

- 視線ヒートマップの分析により，実験参加者間で視線の分布に一定の差異が存在することが確認された．
- 画像の内容によって，実験参加者間の視線パターンの類似性が変化することが明らかになった．
- 相関行列の分析から，全実験参加者ペア間で同程度の差異が観察された．この結果は，特定の個人間に限ら

ず，すべての個人間で識別可能な差異が存在することを示唆しており，個人識別の観点で非常に有望である．

今後の研究の方向性としては，個人の視線パターンに特化した顕著性マップ生成モデルの開発が挙げられる．具体的には，顕著性マップ生成モデルを特定の個人の視線データでファインチューニングすることで，そのユーザが注視する可能性が高い領域を予測することが可能となる．この手法を応用し，ある時点までデバイスを使用していたユーザ A の顕著性マップと，現在のデバイス使用者の実際の注視点分布を比較することで，現在の使用者がユーザ A であるか否かを識別する手法の開発を目指す．これらの取り組みを通じて，より安全で使いやすい認証システムの実現に向けた研究を進めていく予定である．本研究の成果は，ヒューマンコンピュータインタラクション分野における新たな認証手法の基礎となり，セキュリティとユーザビリティの両立に貢献することが期待される．

参考文献

[1] Farzand, H., Abraham, M., Brewster, S., Khamis, M. and Marky, K.: A Systematic Deconstruction of Human-Centric Privacy & Security Threats on Mobile Phones, *International Journal of Human–Computer Interaction*, pp. 1–24 (2024).

[2] Stokkenes, M., Ramachandra, R. and Busch, C.: Biometric Authentication Protocols on Smartphones: An Overview, *Proceedings of the 9th International Conference on Security of Information and Networks*, SIN '16, New York, NY, USA, Association for Computing Machinery, p. 136–140 (online), DOI: 10.1145/2947626.2951962 (2016).

[3] Wei, T. and Zhang, Y.: Fingerprints On Mobile Devices: Abusing And Leaking.

[4] Acar, A., Aksu, H., Uluagac, A. S. and Akkaya, K.: A usable and robust continuous authentication framework using wearables, *IEEE Trans. Mob. Comput.*, Vol. 20, No. 6, pp. 2140–2153 (2021).

[5] Ehatisham-ul Haq, M., Awais Azam, M., Naeem, U., Amin, Y. and Loo, J.: Continuous authentication of smartphone users based on activity pattern recognition using passive mobile sensing, *J. Netw. Comput. Appl.*, Vol. 109, pp. 24–35 (2018).

[6] Wang, H., Zhan, Z., Shan, H., Dai, S., Panoff, M. and Wang, S.: GAZEploit: Remote Keystroke Inference Attack by Gaze Estimation from Avatar Views in VR/MR Devices, *arXiv preprint arXiv:2409.08122* (2024).

[7] Jiang, Y., Leiva, L. A., Rezazadegan Tavakoli, H., R. B. Houssel, P., Kylmälä, J. and Oulasvirta, A.: UEyes: Understanding Visual Saliency across User Interface Types, *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, New York, NY, USA, Association for Computing Machinery, (online), DOI: 10.1145/3544548.3581096 (2023).

[8] Kümmerer, M., Wallis, T. S. A. and Bethge, M.: Saliency Benchmarking Made Easy: Separating Models, Maps and Metrics, *Computer Vision – ECCV 2018* (Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., eds.), Lecture Notes in Computer Science, Springer International Publishing, pp. 798–814.

[9] Bylinskii, Z., Judd, T., Oliva, A., Torralba, A. and Durand, F.: What do different evaluation metrics tell us about saliency models?, *arXiv preprint arXiv:1604.03605* (2016).

[10] Judd, T., Durand, F. and Torralba, A.: A Benchmark of Computational Models of Saliency to Predict Human Fixations, *MIT Technical Report* (2012).

[11] Linardos, A., Kümmerer, M., Press, O. and Bethge, M.: Calibrated prediction in and out-of-domain for state-of-the-art saliency modeling, *CoRR*, Vol. abs/2105.12441 (online), available from ⟨https://arxiv.org/abs/2105.12441⟩ (2021).