

視線計測と物体検出を統合した観察行動の定量評価： 小児看護 AR 教材の前後比較

大浦 綾華¹ 辻野 睦子¹ 伊藤 俊輔¹ 原田 清美² 森本 昌史² 夏川 浩明^{1,a)}

概要：本研究は、小児看護分野における AR 教材の教育効果を客観的に検証することを目的とした。看護学生を対象に、AR を活用した「幼児の病床環境」に関する観察技術のトレーニングを実施し、介入前後で眼鏡型視線計測装置を用いた病床観察課題を行った。計測装置で撮影した一人称映像に対して観察対象を検出する物体検出モデルを構築し、得られたバウンディングボックスと視線座標の時系列情報を統合することで、視線がどの対象に向けられていたか（注視対象）を判定した。さらに、対象別の総注視時間などの指標として定量化する解析パイプラインを整備した。本パイプラインを用いて 14 名のデータを解析し、観察トレーニング前後の観察行動を比較した結果、介入後にはトレーニング前に十分に注意が向いていなかった対象にも視線が配分されるなど、観察の焦点が変化する傾向を定量的に捉えることができた。以上より、定量化が困難な病床観察技能に対し、視線計測と物体検出を統合した解析により教育介入の影響を評価できる可能性が示唆された。

1. はじめに

小児看護では、入院による子どもの反応や環境が子どもに及ぼす影響を予測できる観察力が環境調整には欠かせない。小児看護学実習は看護学生にとって、こうした環境調整技術や観察技術を実際に体験する貴重な機会である。しかし、小児病棟が縮小化される昨今、混合病棟での実習や子どもとかわる期間が短い実習も多く、学内演習の充実が必要不可欠である。

学内演習ではモデル人形や大人の演じる子ども役で代用されリアリティさに欠けることが教育上の課題である。近年、医学教育にバーチャルリアリティ (Virtual Reality :VR) や拡張現実 (Augmented Reality : AR) を用いた教材が取り入れ始め [1]、中でも看護技術教育への AR の活用は、看護技術の疑似体験およびトレーニング、看護技術のスキルやパフォーマンスの評価など、学生の興味関心が高いことが報告されている [2], [3]。我々は、看護学生を対象に AR を活用した「幼児の病床環境」を体験する教材（以下、AR 教材）を開発し [4]、AR 教材を用いた観察トレーニングを小児看護学実習を目前にした 3 年生に実施した。

近年、視線解析技術の発展により、熟練者の視線の動きや注視傾向を客観的に捉えることが可能となった。視線解析を用いた研究には、熟練した観察技術をもつ看護師は重

要な領域をより注視し、患者の顔面部を重点的に注視する傾向があるとの報告がある [5]。しかし、教育的に応用する研究は十分に進んでいるとは言えず、AR 教材で観察トレーニングを行った介入の教育的効果を明らかにしておく必要がある。そこで、本研究では小児看護分野における AR 教材の教育効果を客観的に検証することを目的とした。

2. 実験方法

研究対象者は、A 大学看護学部 3 年生において小児看護学実習直前に観察トレーニングを受講した学生とした。研究対象者には研究の目的や方法、および参加は任意であり、参加および不参加のいずれにおいても学業成績に一切影響がないことを説明したうえで、同意の得られた学生計 14 名を実験参加者（以下、被験者）とした。角膜反射法を用いた眼鏡型視線計測装置（以下、アイトラッカー）Tobii Glasses 3 (Tobii 社製) を使用して視線の計測を行った。被験者には小児用ベッド、乳幼児モデル人形、点滴などの実際の幼児の病床環境を模した空間内で約 1 分間観察してもらい、その間の視線データをサンプリング周波数 100 Hz で計測した。視線データの収集手順は、視線計測を 1 名につき 2 回行い、まず計測の進め方のみ説明を行った状態で 1 回目の視線を計測、1 回目の計測終了後、後述する観察トレーニングを行い、観察トレーニング後に 2 回目の視線を計測した。なお、観察トレーニングの日程に応じて設定したため、データ収集は全員同日ではなく 3 日間にわけて

¹ 大阪成蹊大学

² 京都府立医科大学

^{a)} natsukawa@g.osaka-seikei.ac.jp

実施した。

- (1) 1回目の計測: Session 1
- (2) AR教材による観察トレーニングの実施
- (3) 2回目の計測: Session 2

さらに、一連の計測後には、研究対象者には自分自身の視線の様子についてフィードバックを行った。学生にとって第三者として視線データを見る機会は少ないことから、自分の行動を俯瞰することに興味・関心を抱く学生が多く見られた。その際、1回目と2回目を比較して1回目に見ることができなかった点を2回目に見るよう意識したというコメントが多く聞かれた。なお、本研究は所属機関の倫理審査委員会の承認を得て実施した。

2.1 観察トレーニングの概要

本トレーニングは、幼児の病床環境において臨床判断に必要な観察技能を養うことをねらいとしている。子どもの入院環境で子どもと家族の安全と安心を保障するためには、看護師の訪室毎の短時間で素早く確実な観察力が重要になる。本AR教材は幼児の病床環境をiPad画面に表示させることで、本来の子どもの動き、病床環境にある物品の位置関係、看護師の立ち位置などを視覚的に体験できる。また、ARをした活用した3次元(3D)のCG映像を現実空間に投影することで、看護実践に近い動きでの観察技術を体験できる。本AR教材は図1に示す通り、モデル人形では表現できない幼児の動きを伴う2つのシナリオを用意し、以下の通り実施した。

まず、本トレーニングの特徴は、ARの活用によって同じ場面を繰り返し再現でき、何度でも観察をやり直すことができる点にある。学生は、初めにscenarioの状況設定を把握したうえでAR教材での観察を行い、次にクイズに回答して正答を確認する。続いて「見えていなかった箇所」や「意識的に観察できていなかった箇所」に気づき、再びAR教材での観察を実施する。この一連を繰り返すことにより、看護基礎教育で求められる小児看護の観察技能の基盤が養われると期待し、この後に続く小児看護学実習での看護実践に繋げていくことを目指す。

3. 統合解析パイプライン

計14名の視線計測後、病床を模した実験環境のうち、被験者が何をどのように観察していたかについて視線データより算出するための解析パイプラインを作成する。アイトラッカーはシーンカメラにより一人称映像を記録し、加えてシーンカメラの映像の座標系における視線データを記録している。また、シーンカメラと視線データはどちらも動的に変化するものであり、何を見えていたか導出するためにはシーンカメラに観察対象がいつどこで写っているかを検出し、検出した対象を視線で捉えていたかを判定する必要がある。そのため今回使用した病床環境の対象に特化した



図1 AR教材での観察トレーニングの2つのシナリオ

物体検出のモデルを構築し、視線計測と物体検出の統合解析を行った。以下に、物体検出モデルの構築と評価、視線計測と物体検出の統合解析について記す。

3.1 物体検出モデルの構築

物体検出モデルの構築にあたり、モデルの学習、それに加え再学習と更新を行った。その際に物体検出モデルのRF-DETR[6]を使用した。RF-DETRはDETR系のTransformerベースの物体検出モデルであり、リアルタイム推論を想定した複数のモデルサイズ(例:Nano/Small/Medium)が提供され、カスタムデータへのファインチューニングにも適する。推論結果として、対象クラスと信頼度に加えてバウンディングボックスを出力する。本研究では、各フレームの検出バウンディングボックスと視線座標を時系列に統合し、視線が向けられた観察対象の関心領域(AOI: Area of Interest)を推定するためにRF-DETR(Medium)を使用した。上記モデルに対して、実験環境に使用した小児用ベッドや点滴など物体を撮影した画像やアイトラッカーのシーンカメラの映像を切り出した画像を計1147枚用意した。学習時にはデータ多様性を高めるため、データ拡張を適用した。各学習サンプルに対して、回転(-15°~+15°)、輝度変化(-19%~+19%)、および最大2.5pxのぼかし処理をランダムに適用し、合計4507枚の学習データを準備し、Objects365事前学習重みをチェックポイントとして初期化し、追加学習を行った。また、学習データに用いたクラス(付与したラベルと対応する物)は、「Drugs:薬品(点滴)」、「bed rails:ベッド柵」、「blanket:ブランケット」、「child:乳幼児モデル人形(全身)」、「child-face:乳幼児モデル人形の顔」、「clamp:クレンメ」、「dripping:点滴筒」、「fixtures:ベッド下ストッパー」、「lock:ベッド柵についているレバー」、「nameplate:ネームプレート」、「pump:輸液ポンプ」、「stand:点滴スタンド」、「toys:おもちゃ」の13個である。

3.2 物体検出モデルの評価

構築した RF-DETR モデルの精度は、物体検出モデルにおいて代表的な評価指標である mAP@50, Precision, Recall を用いて評価し、その有効性を明らかにした。本節では、まず mAP (mean Average Precision) の算出における基礎概念として Precision, Recall, および IoU (Intersection over Union) について述べ、続いて AP および mAP の定義について説明する。

まず、Precision とは、モデルが検出した物体のうち、正しく検出された物体の割合を示す指標であり、検出結果の正確性を表す。一方、Recall とは、実際に存在する物体のうち、モデルが正しく検出できた物体の割合を示す指標であり、検出漏れの少なさを評価する指標である。

なお、Precision および Recall の算出式、ならびに本研究において用いた定義を以下に示す。

$$Precision = \frac{TP}{TP + FP} = \frac{TP}{Prediction} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} = \frac{TP}{Ground Truth} \quad (2)$$

ここで、各指標に用いられる用語は以下のとおりである。

- TP (True Positives) : 正しく検出された物体の数
- FP (False Positives) : 実際には存在しない物体を誤って検出した数
- FN (False Negatives) : 実際には存在するが検出されなかった物体の数

次に、IoU について説明する。物体検出においては、モデルが予測したバウンディングボックスが、正解のバウンディングボックスとどの程度一致しているかを定量的に評価する必要がある。IoU は、予測されたバウンディングボックス (Prediction) と、実際の物体位置を示すバウンディングボックス (Ground Truth) との重なり具合を表す指標である。

IoU は、両者が重なっている領域 (Intersection Area) を、それぞれの領域を統合した和集合 (Union Area) で割ることにより算出される。IoU の値は 0 から 1 の範囲を取り、1 に近いほど予測位置が正解に近いことを示す。一般的に、IoU が 0.5 以上である場合に、その検出結果は正しい検出 (TP) として判定されるが、タスクや評価基準に応じてこの閾値は変更される場合がある。

$$IoU = \frac{Intersection Area}{Union Area} \quad (3)$$

この IoU に基づく判定結果を用いて算出される指標が AP (Average Precision) である。AP は、単一クラスに対する物体検出性能を評価する指標であり、モデルが出力した検出結果を信頼度スコアの高い順に並べたうえで、IoU に基づいて各検出を TP または FP として判定する。その後、検出の閾値を段階的に変化させながら Precision および Recall を算出する。

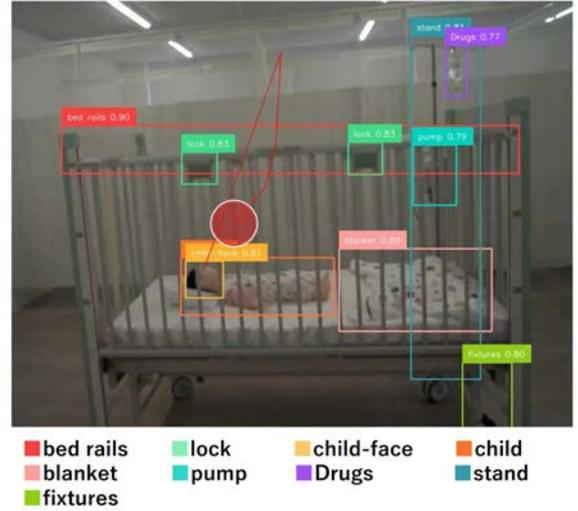


図 2 視線計測と物体検出を統合した後の画像とバウンディングボックスの色とそれに対応する名前 (画像に写っているもののみ記載)

これらの Precision と Recall の関係をプロットすることで、Precision-Recall 曲線 (PR 曲線) が得られる。PR 曲線は、検出閾値の変化に伴う性能推移を示すものであり、テストデータから直接得られた値を結んだ曲線 (Original) と、Precision 値を補間した曲線 (Envelope) によって表される。Envelope は、各 Recall 値に対して、それ以降で得られる最大の Precision 値を用いて構成され、AP の算出にはこの補間後の PR 曲線が用いられる。AP は、この PR 曲線下の面積 (Area Under the Curve : AUC) として定義され、Recall の変動に対してどれだけ安定して高い Precision を維持できているかを示す指標である。

$$AP = \int_0^1 p(r)dr \quad (4)$$

mAP は、すべてのクラスに対して算出された AP の平均値であり、物体検出モデル全体の総合的な性能を評価するために用いられる。例えば、3 クラス (車, 歩行者, 自転車) に対する AP がそれぞれ 0.92, 0.85, 0.78 である場合、mAP はそれらの平均値である 0.85 (85%) となる。mAP は単一の数値でモデル性能を比較できる利点を有しており、物体検出に関する論文やベンチマーク、コンペティションにおいて最も広く用いられている代表的な評価指標である。

$$mAP = \frac{1}{N} \sum AP_i \quad (5)$$

今回は先述した 13 のクラスについての mAP を求めた。またその際の IoU は 0.5 とし、mAP@50 の指標で総合的なモデルの精度を記述した。

3.3 視線計測と物体検出の統合解析

トレーニングにおける被験者の注視行動を分析するため、視線計測データと物体検出結果を統合した解析を行った。統合解析においては、アイトラッカーから取得した注

Total gaze duration on AOI

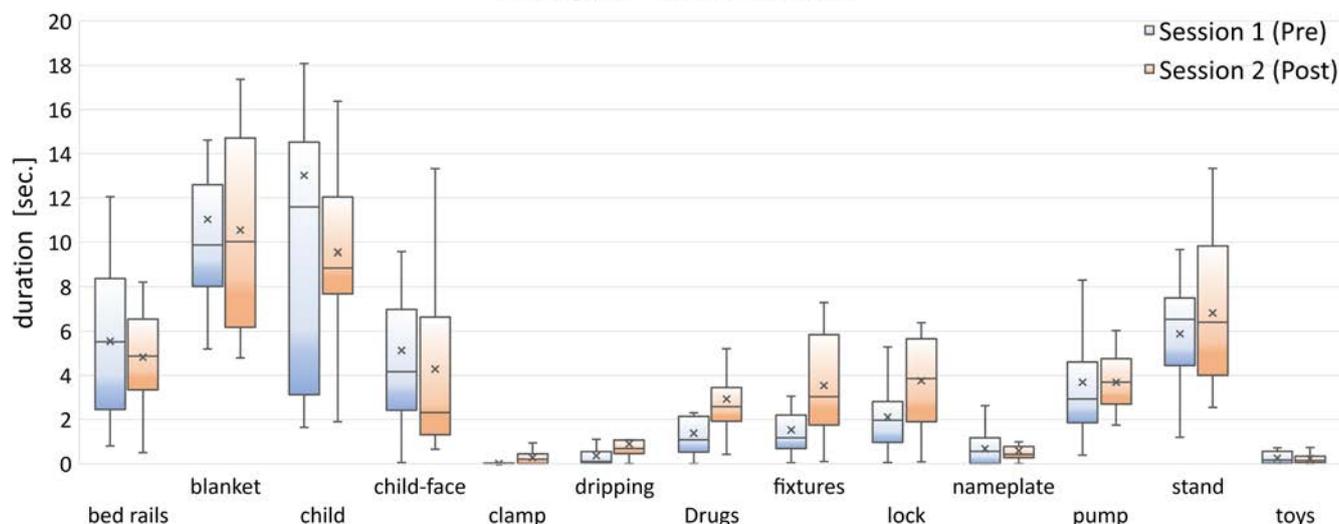


図 3 AR 教材を用いた演習前後のクラス別注視期間 (全被験者) の箱ひげ図

視点の座標情報と、物体モデルによって推定された各物体のバウンディングボックス情報を用い、各視線座標がいずれの物体のバウンディングボックス内に含まれるかに基づいて対応付けを行った。視点によっては物体が重なり、バウンディングボックスが複数重なる場合がある。その際にはバウンディングボックスの中心からの距離が近い物体を見ているものとして扱った。

図 2 に、視線計測と物体検出を統合した後の解析結果の一例を示す。図中では、物体検出によって推定された各物体に対応するバウンディングボックスを色分けして表示しており、合わせて画像内に存在する物体の名称を示している。先の統合解析では画像上においてバウンディングボックス内に赤い点である視線が入っているかどうかを判定したことになる。

また、本研究では観察を始めるために被験者が椅子から立ち上がった瞬間を解析開始時刻とし、椅子に戻ろうとした時間を解析終了時刻として設定した。この区間を対象として解析プログラムを実行することで、観察トレーニングでの学習行動に直接関連する場面における注視行動のみを抽出した。

統合解析の結果として、各物体を観察中にどれほど見ていたかという合計注視時間を物体毎に算出し、AR 教材による観察トレーニング前後である Session 1 と 2 の注視期間の比較を行った。

4. 分析結果・考察

4.1 分析結果について

構築した RF-DETR モデルの精度は mAP@50 が 90.1%、Precision が 85.8%、Recall が 93.1% となり、推論結果を見ても安定的に観察対象を検出できていることを確認した。

また、計測結果の平均値および標準偏差を表 1 に示した。

表 1 計測時間の平均値および標準偏差

計測回	平均計測時間	標準偏差 (SD)
1 回目	01:22:04	00:33:01
2 回目	01:24:56	00:18:50

平均値および標準偏差を算出した結果、1 回目の計測時間の平均値は 1 分 22 秒 (SD=33 秒) であった。一方、2 回目の計測時間の平均値は 1 分 24 秒 (SD=18 秒) であり、平均計測時間の増加と標準偏差の減少が確認された。

視線計測と物体検出の統合解析より図 3 の通り、クラス別注視期間を求め、箱ひげ図で比較した。本結果より、Session 1, 2 ともに乳幼児モデル人形やその顔、ブランケットやベッド柵など、小児患者や比較的大型のものをよく注視している傾向が明らかになった。その中でも Session 1 では乳幼児モデル人形を中心に視覚的に目立つものへの注視が主であったのに対し、Session 2 ではベッド下ストッパーやベッド柵についているレバー、クレンメや点滴筒など、視覚的に目立つもの以外を注視している傾向が見られ、AR 教材を用いた観察トレーニング後に視線行動が変化したことが確認された。

4.2 考察

被験者の計測時間について、表 1 の通り、2 回目の計測では平均計測時間が 1 回目より長くなった一方で、標準偏差が小さくなっていった。このことから、被験者間の計測時間のばらつきが減少し、課題遂行時の行動がより安定していた可能性が示唆される。平均計測時間の増加は、課題内容の理解や注視対象の明確化により、一つ一つの行動に要する時間が増加した可能性があると考えられる。また、標準偏差の減少は、被験者間の個人差が減り、行動パターンが一定の方向に収束したことを示唆する。

「1回目と2回目の計測時間の比較」および「1回目と2回目の全体で見た注視時間の比較」を示した箱ひげ図の結果から、多くの被験者において2回目の計測の方が計測時間および注視時間が長くなる傾向が確認された。特に、2回目の計測では乳幼児モデル人形の全身や顔に向けた視線が減少傾向にあり、観察トレーニングによって被験者の視野が広がり、他に重要視すべき観察するポイントが意識されたと考えられる。このことは、フィードバック時の学生のコメントにも合致している。小児看護学実習前後の学生の視線を比較した先行研究では、病床環境の観察において実習前には視野の中心を注視する傾向が指摘されており [7]、こうした学生の傾向性を踏まえると、本観察トレーニングは、看護学生に小児看護学実習に近いリアルな体験を導き、観察すべき対象を明確にした可能性がある。こうした観察点の明確化によって、視線計測での注視する箇所が増加し、より多くの情報を確認しながら行動していた可能性が示唆された。

今回の視線計測時の環境条件は1回目および2回目の計測において同一であったが、より実際の臨床現場に近い環境を再現するためには、環境条件に複数のパターンを設けることや、計測時間を一定に固定するといった条件を付加した計測を行う必要があると考えられる。これにより、環境要因による影響を制御しつつ、より統制の取れた観察行動の評価が得られる可能性がある。

また、今回の物体検出モデルの精度は mAP@50 で 90% を超え高い精度であったが、クラス別にみるとクレンメや点滴筒など小さい物体の検出精度は他のクラスと比べて低かった。それらのクラスに関しては学習に用いるデータセットの量を増加させることで、モデルの更なる改善ができると考えている。また実際の病床における多様な環境における撮影データなども追加することで、モデルの汎化性能の向上が期待される。

以上の結果を踏まえると、観察トレーニングにおいて観察に要する時間を計測・記録できる仕組みを導入することで、実践前に必要な所要時間を把握できるようになり、観察トレーニングでの学習効果の向上や自己評価の支援につながると考えられる。

5. おわりに

本研究では、小児看護分野における AR 教材の教育効果を客観的に検証した。その結果、本教材での観察トレーニング後には、幼児の病床環境において注視すべき対象が明確になり、より多くの情報を確認しながら行動するという観察の変化が定量化でき、AR 教材を用いた観察トレーニングにより観察技能の向上が示唆された。観察技術を定量的に計測・記録できる仕組みの構築は、小児看護技術習得において実用的であり、今後の研究の継続が看護基礎教育の発展において重要な役割を果たすものと期待される。

謝辞 本研究は JSPS 科研費 24K13686 の助成を受けたものである。また、小児看護 AR 教材の開発にご協力いただいた株式会社スピードの皆様に、ここに記して深謝し上げる。

参考文献

- [1] 小山博史：バーチャルリアリティと拡張現実技術の医学教育／医療応用，日本内科学会雑誌，109(1)．100-106 (2020)．
- [2] Arakida M; Takahashi T; Matsuda Y; Yamabe Y; Miyamoto C; Aoki E; Takahashi S; Ando M; Moriya T; Kodama S; Yamashita R; Igarashi T; Tanizaki K; Uchida K: Investigation of the possibility of using an augmented reality-based endotracheal aspiration simulation tool for nursing education. *Jpn J Nurs Sci.* 21(2), e12573. (2024).
- [3] Chen PJ; Liou WK: The effects of an augmented reality application developed for pediatric first aid training on the knowledge and skill levels of nursing students: An experimental controlled study. *Nurse education today*, 120, 105629 (2023).
- [4] 辻野睦子, 夏川浩明, 伊藤俊輔, 原田清美, 森本昌史：小児看護技術演習での AR 教材の導入を目指した看護学生への予備的調査，*Interaction2025* 発表原稿，4 pages (2025)．
- [5] 上條翔矢, 中山栄純, 小林幹紘, 松谷伸二：患者観察場面における熟練した観察技術をもつ看護師の注視に関するスコーピングレビュー，*日本看護技術学会誌*，22, 51 - 62 (2023)．
- [6] Isaac Robinson, Peter Robicheaux, Matvei Popov, Deva Ramanan, Neehar Peri RF-DETR: Neural Architecture Search for Real-Time Detection Transformers, *arXiv:2511.09554*, (2025)．
- [7] 辻野睦子, 園田悦代：看護学生の学年による小児療養環境における危険認知レベルの相違 眼球運動測定器を用いた検証，*京都府立医科大学看護学科紀要*，26, 19-27 (2016)．