

厚生労働科学研究費補助金（臨床研究等 ICT 基盤構築・人工知能実装研究事業）
『AI の眼』による医療安全確保に関する研究

分担研究報告書

「人工知能技術を用いた手術安全管理システムの社会実装研究」

研究分担者 田淵仁志 広島大学寄付講座教授

【研究要旨】医療安全は人間工学の観点から医療の一領域として分析と研究が行われてきた。特に手術室内での左右、ID、手術用インプラントの取り違い過誤については、多くの報告が上がっており、それらが患者のみならず医療提供者双方に致命的な問題の要因となることは明らかである。ただこれらの防止策として機械的なアプローチは製造業の領域と比較して一般的に行われているとは言い難い。我々は人工知能を用いて手術安全管理を目的とした識別システムを大規模眼科手術実施施設に社会実装するとともに、50%のミスを用意的に混入したストレステストを行うことで、社会実装上の問題点、AIモデル自体の性能について検討を行った。社会実装上では98%以上の認証成功率を示し、ストレステストではAIは100%の認証率を示した。

A. 研究目的

医療安全は人間工学の観点から医療の一領域として分析と研究が行われてきた。ヒューマンエラーは、Reasonによると意図的かどうかによって大別され、されにそれぞれ注意や知識の欠如、悪意などの有無によって細分化され分析されている。それぞれのタイプのエラーについて該当する医療過誤の存在が報告されている。その中の一つである思い込みによる誤認や、同調圧力による事実誤認の問題が古くから指摘されている。チェックする人員を増やせば、逆に誤認率が高まるという研究結果も報告されている。これらのヒューマンエラーに対策として有効とされるのが機械的なチェックシステムであり、医療の現場では投薬管理に多くのシステムが実装されており、製造業では数多くのセンサーシステムが不良品ピックアップのために稼働している。

ID取り違いや左右の間違いなどの医療安全

領域への人工知能モデル応用の取り組みは我々の知る限り報告が無い。我々は、これら3つの医療誤認を抑止するための人工知能技術システムを構築し社会実装に成功している。今回我々はこれら3システムの医療現場での実際の性能を分析するとともに、および50%のミスを用意的に混入したデータセットを用いて行ったAI性能評価ストレステストを施行し研修医との認証率精度の比較検討を行った。

B. 研究方法

社会実装評価

ツカザキ病院（日本、姫路市）眼科手術に施行している顔認証（IDチェック）、左右認証、IOL認証の3種類の人工知能を用いた安全認証システムについて、2021年9月1日から2021年12月28日までの約3か月間に施行された全手術の認証過程の成功率と不成功の要因分析を行った。

ストレステスト

上述の3種類のシステムについて、意図的に50%の不一致データを混在した300個のデータセットに対するストレステストを行った。すべてのテストをAIおよび3名の眼科専門コースに在籍する研修医が行い、その正答率を比較検討した。

社会実装

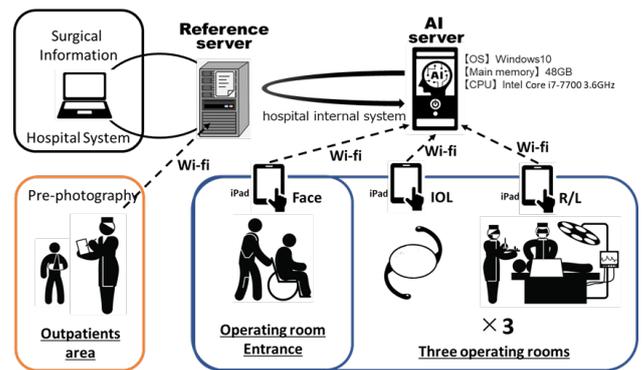
3つのシステムは全てiPad mini(アップル社 アメリカ)のカメラによる撮影写真を用いて認証が行われる。認証用アプリケーションは統合され、アプリケーション内のタブをクリックすることで切り替えられる。撮影された認証用写真はインターネットから隔離された病院内Wifiネットワーク上に置かれたAIサーバに送られ、AIサーバ内でAIモデルを用いて識別同定される。顔認証においては事前に撮影された参照写真との一致度が検討された。左右認証についてはAIモデル内で左右の識別が行われた(図1)。IOL認証についてはAIモデルによる眼内レンズ型番(眼内レンズ名、度数)が識別された後、参照用データベース内に事前登録された挿入用眼内レンズ型番との照合が行われた。手術室は3室あり、iPad miniは各室専用に3台使用された。参照に必要な手術データ(患者ID、術眼、眼内レンズ型番)はインターネットから独立したもう一つの院内ネットワーク上に置かれたデータサーバに置かれており、AIサーバがデータサーバと連携することで認証工程が成立した。

顔認証用の参照用写真事前撮影は眼科外来にて看護師により行われた。手術データ登録は医師により行われた。手術室入室時に看護師により顔認証が行われ、手術室入室後に眼内レンズ認証、および左右認証が行われた。眼内レンズ認証は眼内レンズを

使用する手術でのみ行われた(図2)。3つのAIシステムを用いた安全チェックに加えて、手首バーコードを用いたIDチェック、患者肩部前面にクリップで留められた左右カード、手術直前の看護師、医師3名から5名によるタイムアウトは行われた。さらに眼内レンズ計算、選択および運搬における工程において4名の視能訓練士によるチェックが事前に行われた。

社会実装における認証精度の評価

AIシステムの認証精度についてシステム上に保存された認証過程の記録から検討した。認証ミスについて原因分析を行った。



(図1) 手術安全管理 AI システム構成図

インターネットから独立した院内ネットワークが二つ構成された。安全管理用ネットワークと診療情報格納ネットワークである。参照用サーバ内の正解データとAIサーバ内の手術室内情報を照合することで認証が実施された。手術室は3室あり、それぞれ専用のiPad端末が配置され(計3台)、顔認証、IOL認証、左右認証の3つのアプリケーションは統合アプリケーションとして画面上のタブ切り替えで使用された。顔認証の事前撮影(参照用顔データ)は外来スペースで予め撮影された。全ての認証作業は看護師、視能訓練士によって行われた。



(図 2) ツカザキ病院眼科における AI 手術安全管理システム認証風景

認証は全て看護師によって同一の iPad で行われる。A:外来診察エリアで事前に顔認証参照用写真が撮影されデータベースに登録される。 B : 手術室入口における顔認証風景、C : 手術室内における眼内レンズ認証風景、白いボックスは専用の撮影ボックス、D: 手術開始直前に行う左右認証の風景 覆布がかかった状態で左右認証を行う。覆布開口部と矢印シールを清潔区域から離れた位置から撮影することで認証している。

ストレステスト

顔認証 (IDチェック)

ツカザキ病院眼科にて施行した白内障手術患者 450 人の顔写真を用いた。認証テストは iPad のモニター上で行われた。モニター上に眼科外来で事前撮影された写真 (事前撮影写真) と手術室前室で認証時に撮影された写真 (認証写真) を並べて設置した。2 枚 1 組のテストは 300 セット行われ、そのうち 150 枚が別人のペアになるように並べた。提示順については乱数発生法にて設定された。(図 3)

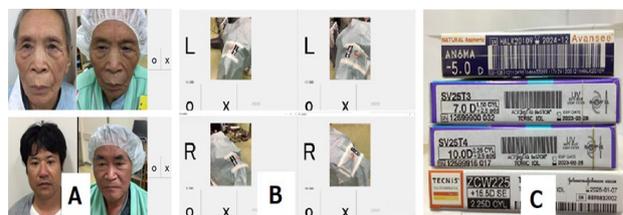
左右認証

ツカザキ病院眼科で行われた白内障手術施行時での左右認証時に取得された認証写真を 300 枚使用した。左右認証写真は覆布開口部と左右認証用矢印シールが明確に判別

できるものであり、写撮影視点方向については頭方向から、左側面から、右側面からの 3 方向が可能な限りランダムであるようにテスト写真をピックアップした。iPad のモニター上に左右認証写真と左右いずれかの記載を提示し、その提示が合っているかどうかについて回答させた。300 枚のうち 150 枚について間違いである方向を記載した。正誤の提示順については乱数発生法にて設定され、10 秒以内での回答を求められた。(図 3)

IOL 認証

未使用眼内レンズ 1200 枚 (1200 種類) を準備した。これらの眼内レンズはツカザキ病院眼科で使用されている IOL チェック用 AI モデルで学習済のものであった。1200 枚を 4 枚セットにして 300 セットを作成した。参照用データベースも同時に作成し、そのうち 150 セットについて間違いデータを登録した。ストレステストは、300 セットが設置された室内で行われた。AI による認証は視能訓練士による iPad 撮影と同時に Wifi ネットワーク上のサーバ上に格納された参照用データに基づいて行われた。研修医による認証は目視により行われた。正誤の提示順については乱数発生法にて設定され、10 秒以内での回答を求められた。(図 3)



(図 3) ストレステスト

A : 顔認証テスト提示画面 左側が事前撮影写真、右側が手術室前室確認写真、B : 左右テスト画面 頭上方向、左側面、右側面の様々な視点方向からの認証写

真が提示された左右表記と一致していれば○を選択し不一致の場合は×を選択する。C：眼内レンズ認証用セットの写真、4枚一組のセットのこの側面と参照用データベースと比較して正誤を判定した。

AI モデル

顔認証は顔認証開発キット ISG-539 (Glory社、日本) を用いて行った。

左右認証モデルは、物体検出モデル (YOLOv3) を 2 段階で用いて、覆布開口部と認証用矢印シールの位置関係を同定し左右を判定した。(第 1 段階) 認証用矢印シールは矢頭部分がヒトの顔に似せてあり、その領域を上方向と同定した (第 2 段階)。上下方向が規定された矢印シールと開口部の位置関係で左右が同定された。第 1 段階の YOLO モデルの学習 (momentum=0.9, decay=0.000001, learning rate=0.001) には 1171 枚の画像を用いた。第 2 段階での学習 (momentum=0.9, decay=0.000001, learning rate=0.001) には 1167 枚の画像を用いた。IOL 認証モデルは 3 段階の AI 認証で構成されており、YOLOv3 による認証対象領域の同定、VGG16[15] による眼内レンズ型番同定、別の VGG16 による眼内レンズ度数同定が行われた。YOLOv3 の学習 (momentum=0.9, decay=0.0005, learning rate=0.001) は 1732 枚で、眼内レンズ型番同定 VGG16 の学習 (momentum=0.001, decay=0.000001, learning rate=0.001) は 30,966 枚 (データ拡張後 123,616 枚 (トレーニング=82,548 枚、バリデーション=41,068 枚))、眼内レンズ度数同定 VGG16 は学習 (momentum=0.001, decay=0.000001, learning rate=0.001) 19317 枚 (データ拡張後 57,126 枚 (トレーニング=45,802 枚、バリデーション=11,324 枚)) で行われた。すべての AI 認

証は Operating System は Windows10 Pro、Main memory は 48GB, CUP は Intel Core i7-7700 3.6GHz という構成のサーバ上で行われた。

C. 研究結果

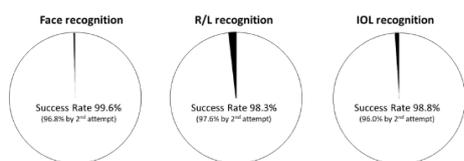
社会実装結果

対象となった手術件数は 3350 件であった。その内訳は眼内レンズ挿入を含む手術が 1874 件、それ以外が 1476 件であった。対象患者の年齢は平均 (SD)74.5(14.0) 歳 Range(最小 4 歳、最大 97 歳)、執刀医の総数は 25 人 (医師ひとり当たりの平均症例数は 134 眼、最小件数は 1 眼、最大件数は 587 眼)、右眼の件数が 1621 眼 (48.3%)、女性が 1752 人 (52.2%)、さらに全身麻酔症例が 97 眼 (2.9%) であった。執刀医の他に麻酔科医 4 名、看護師 31 名、視能訓練士 15 名が少なくとも 1 件以上の手術に参加し、タイムアウト確認作業に携わった。AI 認証工程は看護師が行った。全症例で ID、左右、眼内レンズ挿入ミスは生じなかった。

3350 件の顔認証のうち認証成功率は 3335 件 (99.6%) であった。そのうち必要認証回数は 1 回が 2726 件 (81.4%)、2 回が 514 件 (15.4%)、3 回が 80 件 (2.4%)、4 回が 10 件 (0.3%)、5 回が 5 回 (0.1%)、認証失敗が 15 回 (0.4%) であった。認証失敗症例の全てがマスク着用眼であった。

3350 件の左右認証のうち認証成功率は 3291 件 (98.3%) であった。そのうち必要認証回数は 1 回が 3116 件 (93.0%)、2 回が 154 件 (4.6%)、3 回が 16 件 (0.5%)、5 回が 5 件 (0.2%) であった。認証失敗が 59 回 (1.7%) であり、そのうち撮影距離が遠いことによる撮影不良 13 件、撮影状況不良 (矢印シール貼り忘れ、医師の手や消毒液で開口部が写っていない) が 46 件であった。

1874 件の眼内レンズ認証のうち認証成功率は 1853 件 (98.9%) であった。そのうち必要認証回数は 1 回が 1668 件 (89.0%)、2 回が 131 件(7.0%)、3 回が 42 件 (2.2%)、4 回が 8 件 (0.4%)、5 回が 4 件 (0.2%) であった。認証失敗が 21 回 (1.1 %) であり、そのうち 13 件レンズ未学習であり、4 件が緊急手術による参照用データベース未登録、4 件が撮影不良 (ピンぼけ、レンズの端が切れる) であった。(図 4)



(図 4) ツカザキ病院における AI システム認証成功率

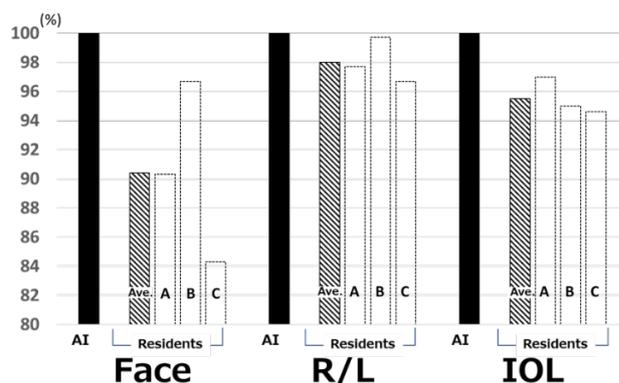
AI による手術安全管理システム社会実装の評価

顔認証 (ID チェック) は認証成功率 99.6% で 96.8%の症例が 2 回目までの認証で成功した。同様に R/L チェックは認証成功率 98.3%、2 回目までの認証で 97.6%成功、IOL チェックは認証成功率 98.8%、2 回目までの認証で 96.0%成功した。

ストレステスト

顔認証 (ID チェック) について AI の正答率は 100%で1問あたりの平均所要時間 2.4 秒であり、研修医 3 名の正答率は 90.3%、96.7%、84.3%で平均(SD)90.4(6.2)% で 1 問あたりの平均(SD)所要時間 3.2(3.3)秒であった。左右認証について AI の正答率は 100%で 1 問あたりの平均所要時間は 1.7 秒であり、研修医 3 名の正答率は 97.7%、99.7%、96.7%で平均 (SD)98.0(1.5)%で1問あたりの平均(SD)所要時間は 2.3(0.5)秒であった。IOL 認証について AI の正答率は 100%で 1 問あたりの平均所要時間 8.2 秒であり、研修医 3 名の正答率

は 97.0 %、95.0 %、94.6 % で平均 (SD)95.5(1.3)%、認証時間平均(SD)値は 10.0 秒であった。(図 5)



(図 5) ストレステスト結果

顔認証 (Face)、左右認証 (R/L)、IOL 認証(IOL)それぞれの AI と研修医 3 名 (A,B,C) とその平均値 (Ave.)(%)

D. 考按

顔認証、左右認証、眼内レンズ認証の 3 つの AI システムの実臨床における認証成功率は全て 98%を超える精度であった。それぞれ 80%以上の症例で 1 回の認証で成功しており、2 回目までの認証成功率は顔認証で 96.7%、左右認証で 97.6%、IOL 認証で 96.1%であった。このことは、総計 75 名の医療従事者が参加する大規模手術室の中で定着し信頼される工程となるために必要な精度と認証効率を兼ね備えていることを示唆している。

さらに、顔認証による ID 識別、矢印シールを用いた左右認証、IOL 認証のストレステストすべてにおいて AI 識別は 100%の精度を示し、研修医 3 名の平均値より良好であった。実装されているシステム、認証テストは全て院内 Wifi ネットワークシステムを用いて行われたが、通信障害等の問題が生じることは無かった。

100%の精度を示す AI モデルであっても臨

床の現場では予め想定していない事象が生じるし、あるいは想定して運用ルールを定めていたとしてもそのルールを逸脱した運用が行われた結果、全体として100%の精度は得られなかった。AI認証作業は自動的にその経過が記録されていくセンサーの役割も果たしており、ヒトによる確認作業が抱える問題点をより鮮明にした。眼内レンズ参照用データ登録作業漏れや撮影不良などの人為的問題混入をいかに減らしていくのかという点は、認証用撮影そのものを機械的に自動的に行うシステムの開発も含めて、実臨床上で今後の重要な課題になるであろう。

ストレステストにおいて研修医3名の認証精度が100%にならなかった事は特別なことではない。どんな単純作業においてもヒトは間違えることは数々の研究で明らかになっている。今回50%のミスという実際の現場ではあり得ない高率のミスの混入によるストレステストによって、ヒトが機械的チェックよりも劣る精度でしか判断できない事が明確に示された事は医療安全学の観点から重要なデータとなった。時間制限があり、しかも件数が多いというストレス下においては、顔認証で10%、左右認証で2%、IOLチェックで5%の確率でミスが混入する可能性があり得るのであるから、そのような環境が実際に生じた際には、現場運営を普段よりもさらに慎重に進める必要がある。今回担当した研修医3名は全て眼科専門医コースに所属する医師である。すなわち眼科医療に関する知識が高く、特にIOLチェックにおいては特にその識別能力が高い集団であることが想定され、看護師や手術アシスタントによるIOLチェック運用が行われている場合、5%以上の誤認率が推定される。

今回の検討の結果が示唆するもう一つ重要な点は、医療安全という課題がAIにとって、最適な課題であることである。ID識別、左右認証、IOL認証いずれも境界例がほぼ存在しない離散値課題である。一方で通常の医学的診断課題はそのほぼすべてに境界例が存在する連続値課題である。課題の特性の差が結果として、医療安全用AIアプリケーションに偽陽性ゼロという日常臨床応用において非常に重要な性能をもたらしたと考察する。事前確率が一般的に小さい医療における識別過程でAIモデルが精度100%でなければ、偽陽性が多発し運用が滞るため、AI利用が回避される結果を招いてしまう。離散値課題を対象とする医療安全AIの識別率100%という性能は、AI医療応用の最大テーマに成り得る可能性がある。さらに、DLが抱えるブラックボックス性を医療安全課題は問題視しなくて良いという面も存在する。これは単純に課題そのものがヒトにとって非常に簡単であることに起因する。左右の識別の際にAIが出す回答について解釈性は求められない。慎重に見れば分かる右か左の判断について、AIの判断プロセスをヒトが要求する必要がないからである。今回のストレステストで用いた症例数がAI性能を評価するに十分であったかどうかは不明である。ただ少なくともこの症例数において眼科専門の研修医3名よりは精度が高いことを示した。さらに顔認証におけるマスク着用課題、最新眼内レンズに対応するための継続的なAI学習追加スキーム確立の必要性など、テクノロジー特有の更新作業もこの方法の課題である。通信ネットワーク構築やモデル管理など技術や費用面についてはチェックリストとは比較にならないぐらい負担が高くなる。ただそれにしてはヒューマンエラーをバックアップする方

法論としての AI 安全管理の優位性は非常に魅力的で AI は同調圧力や思い込みバイアスを持たない独立性があり、ヒトのように集中力の欠如がなく電源がある限り識別能力を維持する。たとえ 1 例でも起こしてはならない医療安全過誤について AI 技術は今後非常に重要な方法論になると考える。

E. 結論

今回我々は識別率 100%を示す AI 医療安全管理システムを開発した。医療の質を支える基幹技術の一つとして今回の手法は今後発展し社会実装されていく可能性がある。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

投稿準備中

2. 学会発表

- 田淵仁志. 人工知能技術 (AI) で良くなる眼科医療. やまぐち眼科フォーラム. 2022/1/15.
- 田淵仁志. 白内障手術と AI アップデート. 第 10 回 JSCRS ウィンターセミナー. 2021/12/15-21. Web 講演
- 田淵仁志. マーケット・イン眼科医療のニーズから AI 応用を考える. 人工知能 (AI) ICT を活用して未来の医療を創る医療人 2030 育成プログラム. 2021/11/27.
- 田淵仁志. 第 1 回: “AI って何? その基本と医療分野への AI の活用事例の紹介”. 第 2 回日本眼科 AI 学会 教育セミナー. 2021/11/4.
- 田淵仁志. 眼手術と AI. 第 75 回日本臨床眼科学会シンポジウム. 2021/10/28-31.

6. 田淵仁志. 近未来の眼科を生み出す人工知能システム. 第 6 回彩の国眼科研究会. 2021/10/23.

7. 出口帆空, 田淵仁志. 人工知能による手術安全システムの左右顔識別実証成績. 第 36 回日本白内障屈折矯正手術学会学術総会. 2021/7/28-8/27.

8. 石飛直史, 田淵仁志. 人工知能による顔認証手法を用いた手術前本人確認システムの性能検証. 第 36 回日本白内障屈折矯正手術学会学術総会. 2021/7/28-8/27.

9. 田淵仁志. 白内障手術と AI. 第 36 回日本白内障屈折矯正手術学会学術総会. 2021/6/26. シンポジウム

10. 田淵仁志. マーケット・インの視点で考える AI 臨床応用. 第 3 回日本メディカル AI 学会. 2021/6/11. Web シンポジウム

11. 田淵仁志. 求められる医療 AI リテラシーとマネジメント. 第 64 回日本糖尿病学会. 2021/5/20-22.

12. 田淵仁志. 臨床応用から考えた AI. 第 6 回メディカル EYE フォーラム IN 沖縄. 2021/4/24.

13. 田淵仁志. AI・日常診察や手術における活用. 日本眼科医会生涯教育講座. 2021/4/18.

14. 田淵仁志. AI と手術室安全管理. 日本区域麻酔学会 第 8 回学術集会. 2021/4/9-10.

H. 知的所有権の取得状況

1. 特許取得

申請検討・準備中

2. 実用新案特許

なし

3. その他

なし