



Aliñamento de imaxes oftalmolóxicas usando representacións neuronais implícitas

Mateo Amado Ares

Dirección: José Rouco Maseda e Jorge Novo Buján

Directores: Jorge Novo Buján
 José Rouco Maseda

Grupo VARPA

Departamento de Ciencias da Computación e Tecnoloxías da Información
Universidade da Coruña





Índice de Contidos

- ① Introdución e Motivación
- ② Contexto e Estado da Arte
- ③ Metodoloxía Proposta
- ④ Experimentos e Resultados
- ⑤ Conclusíons e Traballo Futuro





Tabla de contenidos

- ① Introdución e Motivación
- ② Contexto e Estado da Arte
- ③ Metodoloxía Proposta
- ④ Experimentos e Resultados
- ⑤ Conclusíons e Traballo Futuro

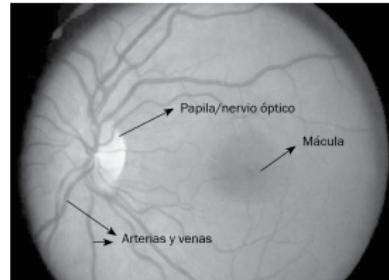
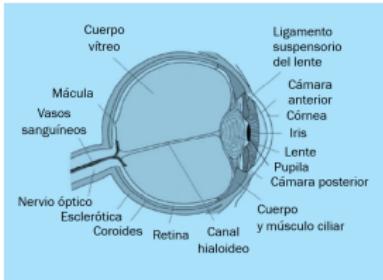




O Problema Clínico: Aliñamento de Imaxes de Retina

Necesidade Clínica

- O aliñamento de imaxes de retina é fundamental para:
 - Rastrear a progresión de enfermidades como o glaucoma ou a retinopatía diabética.
 - Fusionar información de distintas fontes, como retinografía e OCT.
- O ollo permite a observación directa de tecido neuronal e vasos sanguíneos, clave para o diagnóstico precoz.





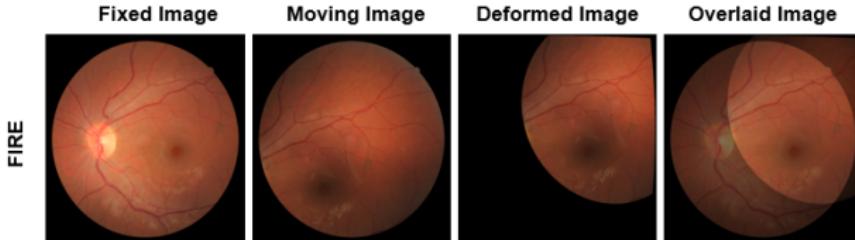
O Reto do Aliñamento Manual

O Reto

O aliñamento manual é un proceso:

- Tedioso e lento: consume tempo dos especialistas.
- Subxectivo e propenso a errores: depende do experto.
- Non escalable: inviable para grandes volumes de datos.

→ A automatización é de gran interese clínico.



Imaxe fixa, móbil, deformada e resultado superposto.





Obxectivos do Traballo

O obxectivo principal é explorar a viabilidade das Representacións Neuronais Implícitas (INRs) para o aliñamento de imaxes oftalmolóxicas.

Obxectivos específicos:

- ① Adaptar o framework IDIR: Modificar a arquitectura orixinal, pensada para imaxes 4D-CT de pulmóns, para rexistro en 2D de retina.
- ② Avaliar o rendemento: Comparar o método en dous conjuntos de datos:
 - FIRE: Imaxes clínicas reais, con variabilidade do mundo real.
 - RFMID: Imaxes sintéticas con transformacións coñecidas.
- ③ Analizar a os resultados: En particular se a activación SIREN ofrece vantaxes para capturar deformacións.





Tabla de contenidos

- ① Introducción e Motivación
- ② Contexto e Estado da Arte
- ③ Metodoloxía Proposta
- ④ Experimentos e Resultados
- ⑤ Conclusíons e Traballo Futuro

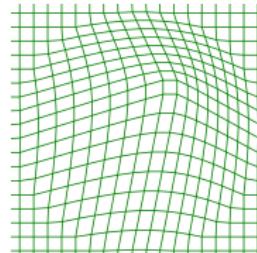
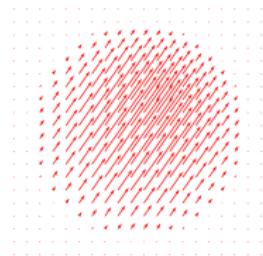


Rexistro Deformable: Campo de Vectores de Deformación

Rexistro Deformable

Atopar unha transformación non ríxida que mapea cada coordenada x da imaxe móbil á súa localización correspondente na imaxe fixa.

- Modélase como un Campo de Vectores de Deformación (DFV), que indica o desprazamento de cada punto.
- Tradicionalmente represéntase o DFV nunha grella discreta de píxeles. Isto ten limitacións de resolución e memoria.





Representacións Neuronais Implícitas (INRs)

Representacións Neuronais Implícitas (INRs)

En lugar dunha grella, representamos a transformación como unha función continua, parametrizada polos pesos dunha rede neuronal (MLP).

- Entrada: Coordenada (x, y) .
- Saída: Vector de desprazamento (dx, dy) .
- Vantaxes clave:
 - Independencia da resolución: A función é continua, non depende do tamaño da imaxe.
 - Gradiéntes analíticos: Permite calcular derivadas exactas da deformación, crucial para unha regularización precisa.



O Problema do Sesgo Espectral

Limitación das Redes con ReLU

As funcións de activación lineais coma ReLU presentan un sesgo espectral cara funcións de baixa frecuencia (transformacións globais e ríxidas), e teñen dificultades para representar detalles finos e cambios locais.

No contexto do rexistro de retinas, isto é problemático porque os detalles de alta frecuencia son críticos para un aliñamento preciso.



SIREN (Sinusoidal Representation Networks)

A Solución:

SIREN utiliza a función seno como activación para superar o sesgo espectral e modelar detalles finos.

$$f(x) = \sin(ax + b), \quad \text{con } a, b \in \mathbb{R}$$

- Ideal para sinais complexos: Permite representar detalles de alta frecuencia e as súas derivadas.
- Infinitamente diferenciable: A diferencia de ReLU, a función seno pode derivarse cantas veces sexa necesario, permitindo regularizacíons de orde superior.





Estado da Arte: Métodos Clásicos

Rexistro baseado en características (FBR)

Identifican, describen e emparellan características salientables entre imaxes.

- Robustos ante variacións locais e cambios de modalidade.

Rexistro baseado en intensidade (IBR)

Comparan directamente os valores de intensidade dos píxeles mediante unha métrica de similitude e optimizan iterativamente a transformación.

- Non requieren a extracción de características explícitas.
- Máis sensibles a ruido e variacións de iluminación ou estrutura.





Estado da Arte: Métodos de Aprendizaxe Profunda

Dúas Estratexias Principais

As redes neuronais abordan o rexistro de dúas formas:

- Aprendizaxe Extremo a Extremo (End-to-End): Predí directamente a deformación a partir dun par de imaxes.
- Mellora de Métodos Clásicos: Empréganse redes para substituír compoñentes das metodoloxías clásicas.

O noso enfoque: Modelo de Transformación

O noso método, IDIR, aprende unha representación implícita da transformación parametrizada nos pesos da rede neuronal para cada par de imaxes.





Estado da Arte en Retinografías

Desafíos

Rexistro difícil por deformacións non ríxidas, baixo solapamento e variacións de iluminación e anatomía.

Solucións Actuais e Contribución

- Tendencia Actual: Métodos Híbridos: Combinan deep learning e algoritmos clásicos para lograr precisión e velocidad.
- A Nosa Contribución: Primeira aplicación de Representacións Neuronais Implícitas (INRs) ao rexistro de retinografías, adaptando o framework IDIR a este dominio.





Tabla de contenidos

- ① Introducción e Motivación
- ② Contexto e Estado da Arte
- ③ Metodoloxía Proposta
- ④ Experimentos e Resultados
- ⑤ Conclusíons e Traballo Futuro

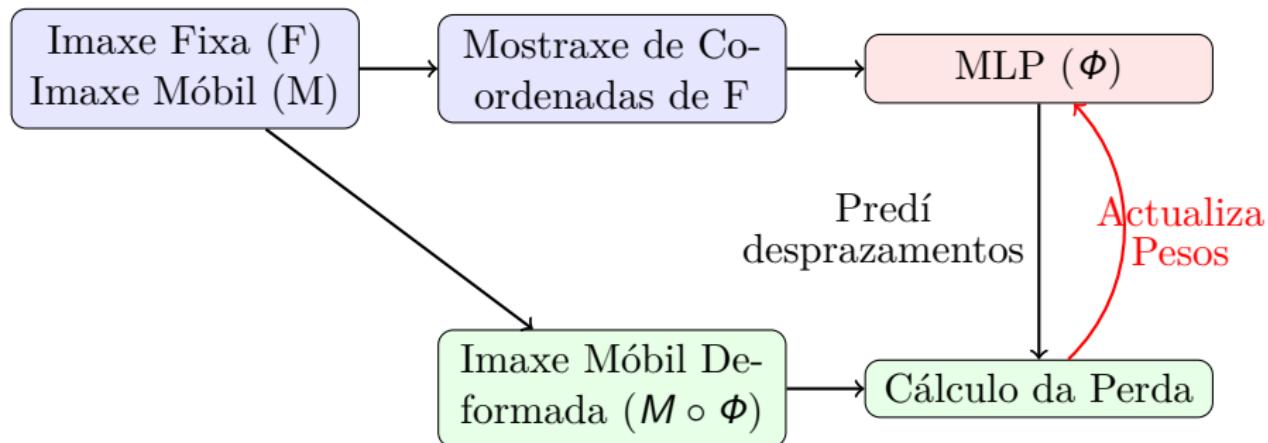




Framework IDIR: Esquema do Proceso

Optimización por Par de Imaxes

A metodoloxía de IDIR adestra unha nova rede para cada par de imaxes (Fixa F e Móbil M).





Framework IDIR: Función de Perda

Función de Perda

A rede optimízase para minimizar unha perda combinada:

$$\hat{\phi} = \operatorname{Arg} \min_{\phi} L_{data}(M \circ \Phi, F) + \alpha L_{reg}(\Phi)$$

- \mathcal{L}_{datos} : Mide a similitude entre a imaxe fixa e a móbil deformada (e.g., NCC).
- \mathcal{L}_{reg} : Penaliza as deformacións non realistas para garantir a suavidade.

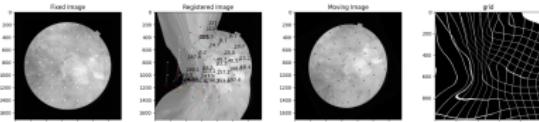


Framework IDIR: Regularización

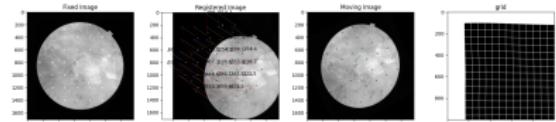
Problema:

O rexistro de imaxes é un "ill-posed problem": existen moitas deformacións que poden facer que as imaxes se parezan, pero a maioría non son solucións correctas (non fisicamente realistas).

A regularización engade coñecemento previo físico para restrinxir o espazo de solucións posibles.



Sen regularización: pregamentos non físicos.



Exceso de regularización: deformación demasiado ríxida.



Framework IDIR: Regularizador Xacobiano

- Controla a distorsión local da área e prevé pregamentos (folding) non realistas na deformación, evitando compresións ou estiramentos extremos.
- Mecanismo: Penaliza as desviacións do determinante da matriz Xacobiana ($\det \nabla \Phi$) respecto do valor 1.
 - Un determinante próximo a 1 preserva a área local.
 - Un determinante negativo ou cero indica un pregamento da malla, o que invalida a transformación.

$$S^{\text{jac}}[\Phi] = \int_{\Omega} |1 - \det(\nabla \Phi)| \, dx \, dy$$





Framework IDIR: Regularizador Hiperelástico

- Mecanismo: Combina termos que penalizan a variación da lonxitude dos vectores de desprazamento e as distorsións de área.
 - O termo $|\nabla u|^2$ controla o estiramento.
 - O termo co cofactor (cof) controla a deformación da área.
- É un regularizador máis xeral que engloba propiedades do Xacobiano e impón unha certa suavidade.

$$S^{\text{hyper}}[\phi] = \int_{\Omega} \left[\frac{1}{2} \alpha_I |\nabla u|^2 + \alpha_a \varphi_c(\text{cof} \nabla \phi) \right] dx dy$$

con

$$\varphi_c(C) = \sum_{i=1}^2 \max \left\{ \sum_{j=1}^2 C_{ji}^2 - 1, 0 \right\}^2$$





Framework IDIR: Regularizador de Enerxía de Flexión

- Evita cambios bruscos na curvatura do campo de desprazamento, o que se traduce en transformacións visualmente más suaves e continuas.
- Penaliza a magnitude das segundas derivadas parciais do campo de deformación (ϕ) en todo o dominio.

$$S^{\text{bending}}[\phi] = \frac{1}{8} \int_{\Omega} \left[\left(\frac{\partial^2 \phi}{\partial x^2} \right)^2 + \left(\frac{\partial^2 \phi}{\partial y^2} \right)^2 + 2 \left(\frac{\partial^2 \phi}{\partial x \partial y} \right)^2 \right] dx dy$$





Consideracións sobre a Regularización

- Balance da regularización: É crucial atopar un equilibrio no peso dos termos de regularización.
- Custo computacional: A regularización incrementa significativamente o tempo de cómputo, xa que require calcular gradientes adicionais por época.

Vantaxe SIREN

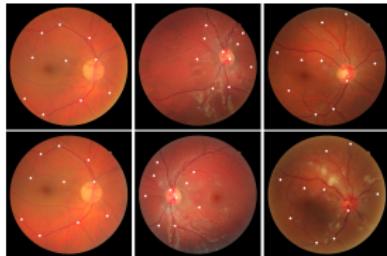
O cálculo de regularizadores de orde superior (como a Enerxía de Flexión) de forma eficiente e analítica só é posible porque as redes SIREN son infinitamente diferenciables. Arquitecturas baseadas en ReLU non poden usar este termo, xa que a súa segunda derivada é cero en case todos os puntos.



Datasets Utilizados

FIRE

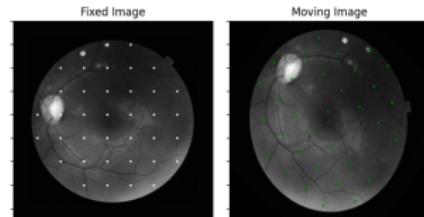
- 134 pares de imaxes clínicas reais.
- Diferentes graos de variacións en iluminación, patoloxías e superposición.



FIRE

RFMID

- Xeramos pares de imaxes aplicando transformacións xeométricas coñecidas e controladas.
- Sen variacións de apariencia.

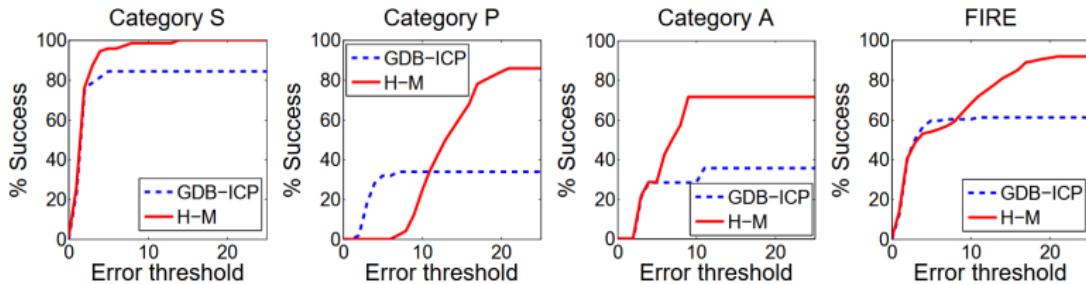


RFMID

Métrica de Avaliación

Protocolo de Avaliación

Usamos o protocolo estándar de FIRE: unha gráfica que mostra a porcentaxe de rexistros exitosos para un límiar de erro crecente. Máis alto e á esquerda é mellor.

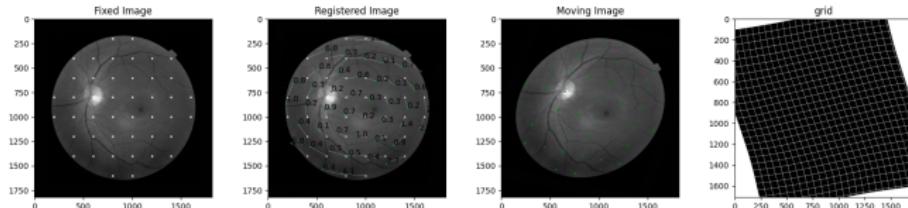


Análise Cualitativa

Análise Visual

A avaliación visual é crucial para entender como e por que o método funciona ou falla.

Rexistro Exitoso (RFMiD, ReLU)



A grella de deformación é suave e os puntos de control (azul) axústanse correctamente.



Tabla de contenidos

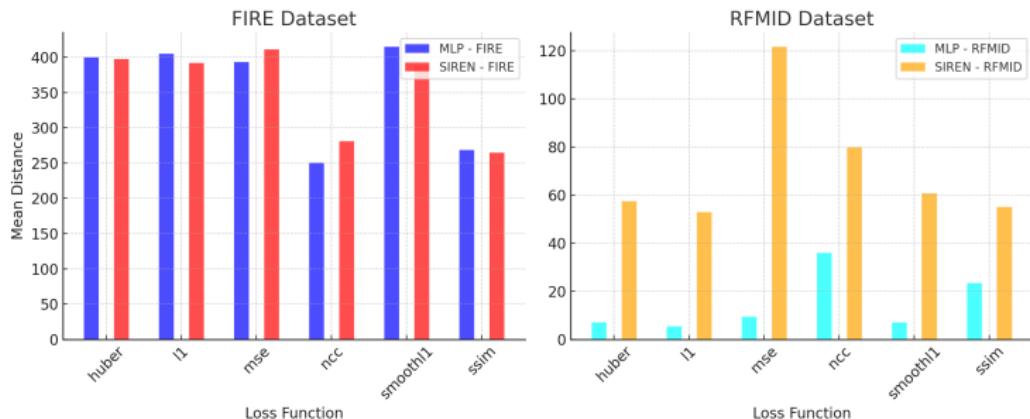
- ① Introducción e Motivación
- ② Contexto e Estado da Arte
- ③ Metodoloxía Proposta
- ④ Experimentos e Resultados
- ⑤ Conclusíons e Traballo Futuro



Experimentos: Función de Perda

Motivación

Atopar a métrica de similitude máis robusta para as imaxes de retina.

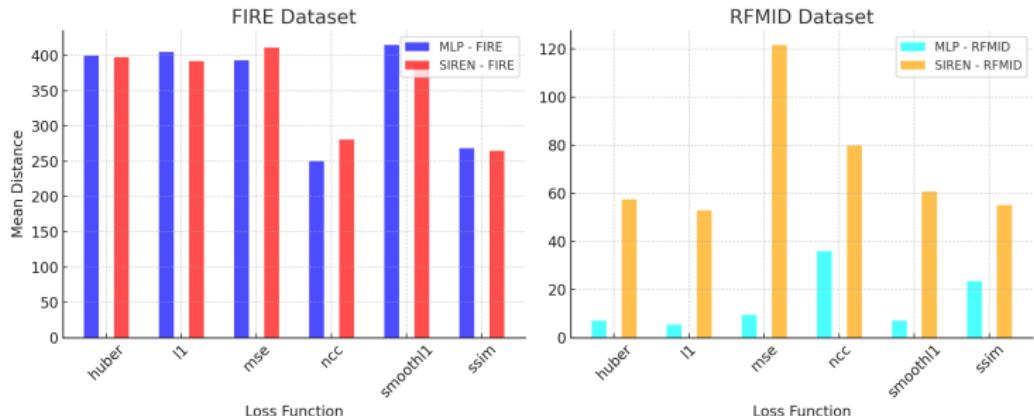


Experimento: Función de Perda

Conclusión

A métrica ideal depende da natureza dos datos:

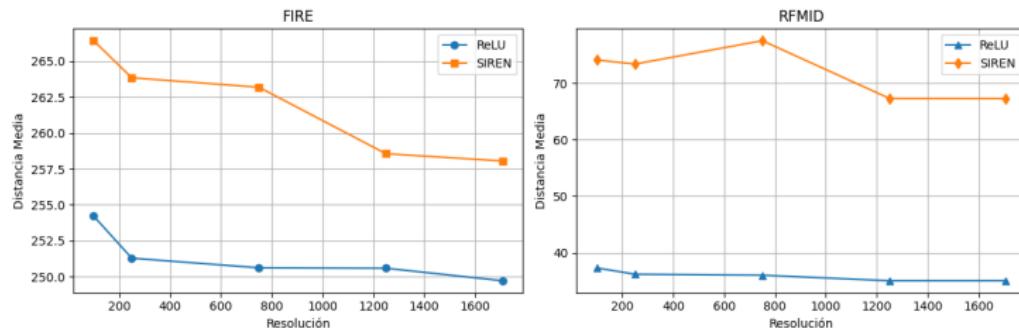
- FIRE: Métricas estruturais como NCC funcionan mellor.
- RFMID: Métricas baseadas en píxeles como L1 ou MSE son superiores.



Experimento: Resolución da Imaxe

Motivación

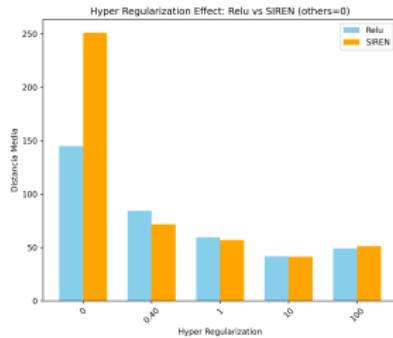
Determinar se usar imaxes de alta resolución (ata 2160x2160px) mellora o resultado ou só aumenta o custo computacional.



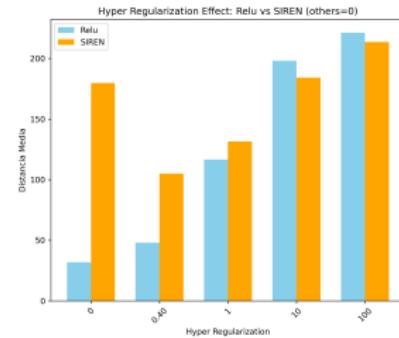
Conclusión

Non se observa unha mellora significativa por enriba de 1250x1250 píxeles. Este tamaño ofrece un bo equilibrio.

Experimento: Regularización



Regularización
hiperelástica en FIRE

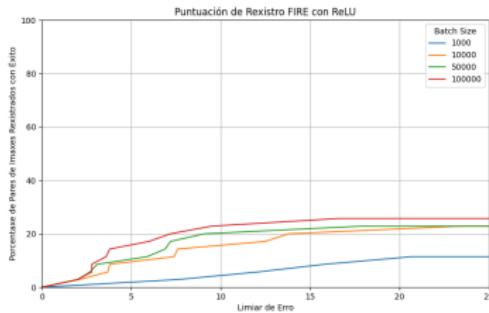


Regularización
hiperelástica en RFMID

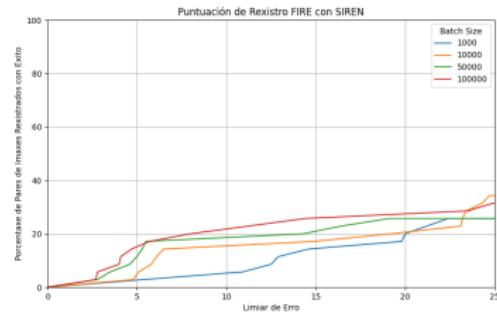
Conclusión Principal

A regularización é indispensable para SIREN. A cantidade óptima depende transformación concreta: deformacións más complexas (FIRE) benefíxianse de maior regularización.

Experimento: Tamaño do Lote



ReLU (FIRE - Cat. S)



SIREN (FIRE - Cat. S)

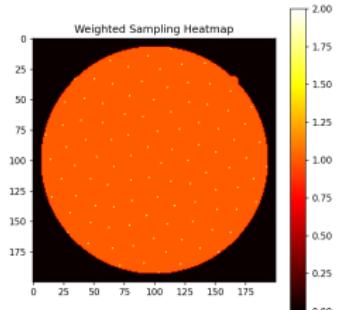
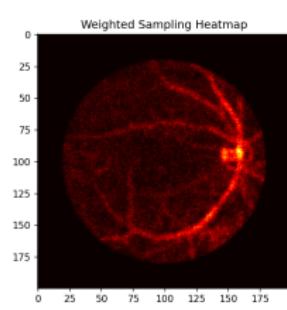
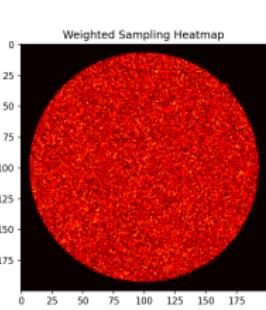
Conclusión

Un tamaño de lote $>10,000$ é fundamental para obter bons resultados. Tamaños moi superiores non ofrecen beneficios significativos.

Estratexia Desenvolvida: Mostraxe Intelixente (I)

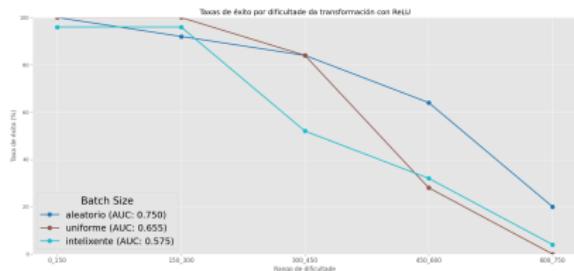
Hipótese

Priorizar rexións con máis información estrutural (vasos sanguíneos, disco óptico) mediante estratexias de mostraxe "intelixentes" podería mellorar o resultado final do rexistro fronte á mostraxe puramente aleatoria.

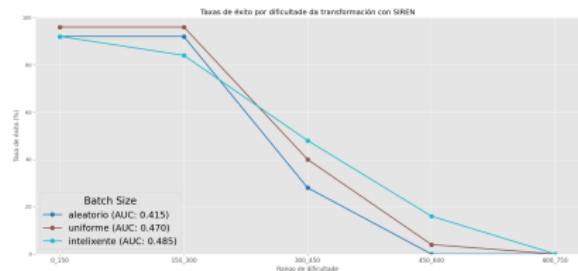


Estratexia Desenvolvida: Mostraxe Intelixente (II)

ReLU (RFMID)



SIREN (RFMID)



Resultado e Conclusóns

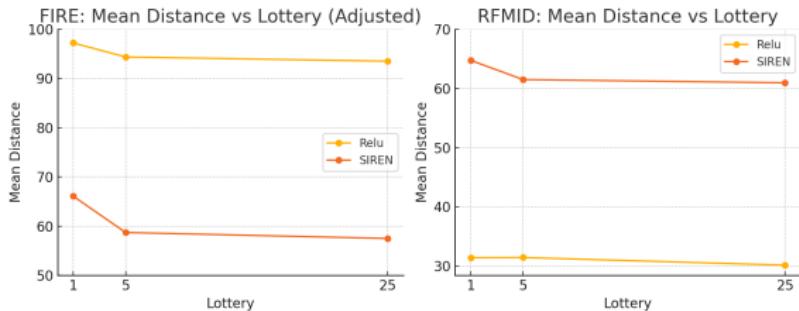
En contra da hipótese, ningunha das estratexias de mostraxe implementadas ofrece unha vantaxe significativa. Isto indica que a información relevante está suficientemente dispersa, sendo a mostraxe aleatoria axeitada sempre que o tamaño do lote sexa grande dabondo.



Estratexia Desenvolvida: Lotería de Inicialización (I)

Hipótese

Explorar varias inicializacíons de pesos aleatorias e escoller a máis prometedora (baseándose na perda inicial) podería evitar mínimos locais deficientes e mellorar o rendemento final.

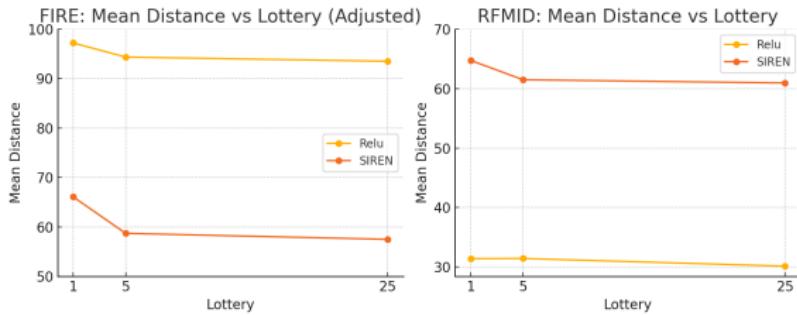




Estratexia Desenvolvida: Lotería de Inicialización (II)

Resultado e Conclusión

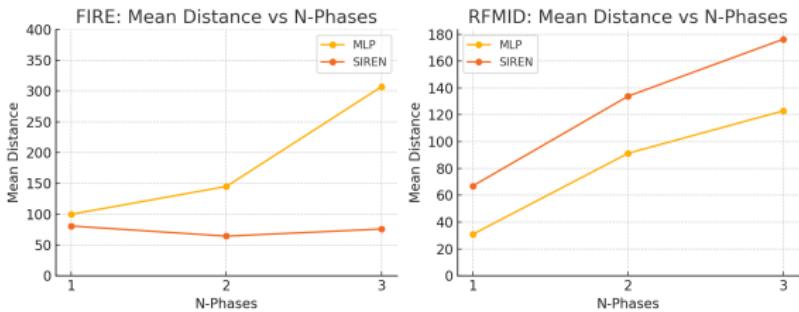
Obsérvase unha mellora marxinal no rendemento. A inicialización ten certo impacto, mais o seu custo computacional adicional pode non xustificar a pequena ganancia.



Estratexia Desenvolvida: Axuste Dinámico do Lote (I)

Hipótese

A rede podería beneficiarse dun adestramento por fases: comenzar cun tamaño de lote pequeno para aprender a transformación global e despois aumentalo progresivamente para refinar os detalles locais da deformación.



Estratexia Desenvolvida: Axuste Dinámico do Lote (II)

Resultado e Conclusión

O resultado foi concluínte e contrario á hipótese: esta estratexia resultou ser prexudicial, empeorando o rendemento a medida que se incrementaba o número de fases. Un tamaño de lote grande e constante dende o inicio demostrou ser máis estable e eficaz.

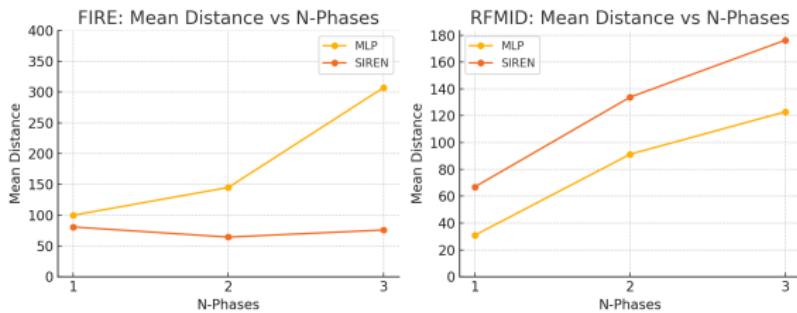




Tabla de contenidos

- ① Introducción e Motivación
- ② Contexto e Estado da Arte
- ③ Metodoloxía Proposta
- ④ Experimentos e Resultados
- ⑤ Conclusíons e Traballo Futuro





Conclusións

Resumo

O framework IDIR funciona para rexistro de retinografías, pero a súa dependencia só da intensidade dos píxeles limita a robustez clínica.

- ReLU: boa en transformacións globais e lineais (RFMID).
- SIREN: eficaz en deformacións locais (FIRE), pero require regularización más coidadosa.
- Limitación clave: Falta de correspondencia global; o método é fráxil ante grandes desprazamentos, variacións de apariencia e baixa superposición.





Traballo Futuro (I)

Diagnóstico do Problema

As nosas conclusións indican que os modelos INR son bós para o refinamento local de deformacións, pero fráxiles para atopar a correspondencia global inicial.

Solución Proposta: Enfoque Híbrido

Sistema de dous pasos que combina o mellor dos métodos baseados en características e os baseados en aprendizaxe.

- Paso 1: Aliñamento global cun método baseado en puntos clave.
- Paso 2: Refinamento local usando o noso modelo INR para axustar detalles finos.





Traballo Futuro (II): Mellora da INR

Melloras na arquitectura e regularización

- Consistencia Cíclica: Adestrar dúas redes (directa e inversa) que se regularicen mutuamente.
 - Vantaxe clave: Xera unha métrica de certeza da transformación, útil en aplicacións clínicas.
- Garantir a Invertibilidade: Empregar arquitecturas baseadas en campos de velocidade ou redes invertibles para obter transformacións más robustas e realistas.





Traballo Futuro (III): Coñecemento Adicional

Incorporación de coñecemento adicional

- Meta-aprendizaxe: Aprender unha inicialización de pesos óptima a partir dun conxunto de datos para acelerar e estabilizar a converxencia en novas imaxes.
- Condicionamento por xeometría: Incorporar coñecemento anatómico previo (ex: segmentación de vasos) para restrinxir o espazo de posibles deformacións.

Melloras no preprocessado

- Explorar técnicas de preprocessado de imaxes para mellorar a calidade dos datos de entrada e facilitar o rexistro.





Grazas pola atención.





Aliñamento de imaxes oftalmolóxicas usando representacións neuronais implícitas

Mateo Amado Ares

Dirección: José Rouco Maseda e Jorge Novo Buján

Directores: Jorge Novo Buján
 José Rouco Maseda

Grupo VARPA

Departamento de Ciencias da Computación e Tecnoloxías da Información
Universidade da Coruña

