

MÉTODO DE PROCESAMIENTO DE AL MENOS UNA IMAGEN DE PRECONTRASTE Y UNA IMAGEN DE CONTRASTE QUE REPRESENTAN RESPECTIVAMENTE UNA PARTE DEL CUERPO ANTES Y DESPUÉS DE UNA INYECCIÓN DE UNA PRIMERA DOSIS DE AGENTE DE CONTRASTE

CAMPO DE LA INVENCION

El campo de la presente invención es el del aprendizaje automático/profundo.

Más particularmente, la invención se refiere a métodos de procesamiento de al menos una imagen de precontraste y una imagen de contraste que representan respectivamente una parte del cuerpo antes y después de una inyección de una primera dosis habitual de agente de contraste, en particular para generar una imagen de contraste sintética que simula dicha parte del cuerpo después de una inyección de una segunda dosis de agente de contraste que es superior a la dosis habitual.

ANTECEDENTES DE LA INVENCION

Los agentes de contraste son sustancias usadas para aumentar el contraste de estructuras o líquidos dentro del cuerpo en imágenes médicas.

Por lo general, absorben o alteran las radiaciones externas emitidas por el dispositivo de obtención de imágenes médicas. En rayos X, los agentes de contraste realzan la radiodensidad en un tejido o estructura diana. En las IRM, los agentes de contraste modifican los tiempos de relajación de los núcleos en los tejidos corporales para alterar el contraste de la imagen.

Los agentes de contraste se usan habitualmente para mejorar la visibilidad de lesiones, sobre todo en neuroimagen para el diagnóstico inicial y la planificación del tratamiento de tumores cerebrales, tales como gliomas, metástasis cerebrales, meningiomas.

El contraste de susceptibilidad dinámica (DSC) y el contraste dinámico realizado (DCE), que aprovechan respectivamente los efectos T2 y T1, son las dos técnicas más comunes para la IRM. En ambos casos, se inyecta por vía intravenosa al paciente un agente de contraste basado en gadolinio (GBCA) y se realizan imágenes rápidas repetidas para obtener una secuencia temporal de imágenes.

La inyección de GBCA aumenta la sensibilidad de la IRM, lo que permite, por ejemplo, la identificación de nódulos carcinogénicos más pequeños, iniciar antes el tratamiento y, a su vez, mejorar la supervivencia y la calidad de vida de los pacientes. Una solución para

aumentar aún más la sensibilidad de la IRM es aumentar la cantidad inyectada de GBCA.

Sin embargo, basándose en consideraciones de precaución, las directrices clínicas recientes sugieren el uso de la dosis mínima que consiga un realce suficiente del contraste, por lo que el uso de GBCA debe ser lo más parco posible.

Se pueden identificar tres vías de investigación complementarias para mejorar, no obstante, el contraste de la IRM y su rendimiento en la detección de lesiones:

- En primer lugar, los nuevos agentes de contraste con propiedades químicas mejoradas tienen el potencial de mejorar el contraste de la imagen sin aumentar la dosis inyectada de gadolinio. Sin embargo, dichos agentes aún están en fase de prueba y es probable que sean aún más caros que el GBCA.

- En segundo lugar, las novedosas secuencias de IRM, con o sin agente de contraste, pueden sustituir o complementar respectivamente a las secuencias T1 rutinarias realizadas con contraste. En particular, la secuencia turbo espín eco T1 con ángulos de cambio de orientación variables (TSE) demostró ser más sensible que su homóloga de eco de gradiente (GRE) y ahora se recomienda para la obtención de imágenes de tumores cerebrales, aunque también se identificaron algunas limitaciones cualitativas, además del mayor tiempo mínimo de exploración que puede favorecer los artefactos de movimiento.

- En tercer lugar, se están proponiendo cada vez más enfoques prometedores de aprendizaje profundo para detectar y delinear tumores automáticamente, con la promesa de niveles de precisión uniformes y constantes con tiempos de lectura prácticamente instantáneos. Sin embargo, estos algoritmos de aprendizaje profundo aún están en fase de investigación y su rendimiento sigue siendo insuficiente para un despliegue inmediato a gran escala.

En consecuencia, sigue existiendo la necesidad de un nuevo método para mejorar aún más la sensibilidad de las imágenes médicas de contraste realizado.

RESUMEN DE LA INVENCION

A estos efectos, la presente invención proporciona un método de obtención de imágenes médicas, caracterizándose el método por que comprende la implementación, por un procesador de datos de un segundo servidor, de etapas de:

- (a) Obtener (i) al menos una imagen candidata de precontraste y una imagen candidata de contraste que representan respectivamente una parte del cuerpo antes y

después de una inyección de una primera dosis de agente de contraste, en donde la primera dosis de agente de contraste es superior a una dosis baja de agente de contraste y la dosis baja es inferior a una dosis habitual de agente de contraste; y (ii) una red neuronal convolucional, CNN, que está entrenada para reconstruir, a partir de al menos una imagen de entrada de precontraste y una imagen de entrada de contraste de dosis baja que representan respectivamente la parte del cuerpo antes y después de una inyección de la dosis baja de agente de contraste, una imagen de contraste de dosis habitual que representa la parte del cuerpo después de una inyección de la dosis habitual de agente de contraste;

(b) generar una imagen de contraste sintética aplicando la CNN a la al menos una imagen candidata de precontraste y a la imagen candidata de contraste.

Las características preferidas, pero no limitantes, de la presente invención son como sigue:

Dicha dosis baja es entre el 10 y el 50 % de la dosis habitual.

Dicha dosis baja es entre $1/5$ y $1/3$ de la dosis habitual, preferiblemente aproximadamente del 25 %.

Dicha primera dosis es al menos el 50 % de la dosis habitual.

Dicha primera dosis es al menos el 80 % de la dosis habitual, preferiblemente aproximadamente el 100 % de la dosis habitual.

De modo que dicha segunda dosis sea al menos el 240 % de la dosis habitual, preferiblemente entre el 320 % y el 400 % de la dosis habitual.

La al menos una imagen candidata de precontraste y la imagen candidata de contraste son adquiridas por un dispositivo de imagen médica conectado al segundo servidor, en particular un escáner de IRM.

Dicha al menos una imagen candidata de precontraste incluye una imagen de precontraste ponderada en T1 y la imagen candidata de contraste es una imagen ponderada en T1.

Dicha al menos una imagen de precontraste comprende tres imágenes candidatas de precontraste que incluyen la imagen de precontraste ponderada en T1, una imagen de precontraste ponderada en T2-flair y un mapa de precontraste de ADC.

La CNN se entrena para reconstruir, a partir de una imagen de entrada de precontraste ponderada en T1, una imagen de entrada de precontraste ponderada en T2-flair,

un mapa de entrada de precontraste de ADC y una imagen de entrada de contraste ponderada en T1 que representan respectivamente una parte del cuerpo antes y después de una inyección de dicha dosis baja de agente de contraste, una imagen de contraste de dosis baja ponderada en T1 que representa dicha parte del cuerpo después de una inyección de dicha dosis habitual de agente de contraste.

La etapa (b) comprende aplicar la CNN a las tres imágenes candidatas de precontraste y a la imagen candidata de contraste, a fin de generar una imagen de contraste sintética ponderada en T1.

Dicha CNN comprende una rama codificadora seguida de una rama decodificadora, con conexiones de salto entre la rama codificadora y la rama decodificadora.

El método comprende una etapa previa de entrenamiento, por un procesador de datos de un primer servidor, de dicha CNN a partir de una base de secuencias de al menos una imagen de precontraste de entrenamiento, una primera imagen de contraste de entrenamiento y una segunda imagen de entrenamiento que representan respectivamente una parte del cuerpo antes, después de una inyección de dicha dosis baja de agente de contraste, y después de una inyección de dicha dosis habitual de agente de contraste.

Según un segundo y un tercer aspecto, la invención proporciona un producto de programa informático que comprende instrucciones de código para ejecutar un método para la formación de imágenes médicas según el primer aspecto; y un medio legible por ordenador, en el que está almacenado un producto de programa informático que comprende instrucciones de código para ejecutar para la formación de imágenes médicas de dicho método según el primer aspecto.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Los anteriores y otros objetos, características y ventajas de la presente invención serán evidentes en la siguiente descripción detallada de una realización ilustrativa de la misma, que debe leerse en relación con los dibujos adjuntos en donde:

- la figura 1 ilustra un ejemplo de arquitectura en la que se realiza el método según la invención;
- la figura 2 representa un ejemplo de canalización de entrenamiento y canalización de operación de una CNN en el método según la invención;
- la figura 3 representa una arquitectura preferida de CNN usada en el método

según la invención;

- la figura 4 ilustra una realización del método según la invención;
- la figura 5 compara imágenes de contraste T1c y T1c+ obtenidas usando el método según la invención.

DESCRIPCIÓN DETALLADA DE UNA REALIZACIÓN PREFERIDA

Arquitectura

La presente invención propone un método de obtención de imágenes médicas, en particular de procesamiento mediante una red neuronal convolucional, CNN, al menos una imagen candidata de precontraste y una imagen candidata de contraste que representan respectivamente una parte del cuerpo antes y durante o después de una inyección de una primera dosis de agente de contraste.

Por imagen de precontraste, o imagen “plana”, se entiende una imagen que representa una parte dada del cuerpo (que se va a monitorizar) antes de la inyección del agente de contraste, para una persona o un animal. Por imagen de contraste se entiende una imagen que representa dicha parte del cuerpo durante o después de la inyección del agente de contraste.

En otras palabras, si hay una secuencia temporal de imágenes, la primera es la imagen de precontraste, y cada una de las siguientes es una imagen de contraste. Obsérvese que las imágenes de contraste pueden ser imágenes de una fase dada (por ejemplo, arterial, portal, retardada) o contraste totalmente dinámico realzado (DCE).

En la siguiente descripción, los presentes inventores tomarán el ejemplo preferido de imágenes cerebrales, es decir, dicha parte del cuerpo es el cerebro.

Las imágenes (de precontraste o contraste) son adquiridas directamente, o derivadas de imágenes adquiridas directamente, por un dispositivo de imagen médica del tipo escáner.

Dichas imágenes con inyección de agente de contraste pueden ser:

- TC (tomografía computerizada) → el dispositivo médico de formación de imágenes es un escáner rotacional de rayos X capaz de reconstrucción tomográfica;
- IRM (imagen por resonancia magnética) → el dispositivo de formación de imágenes médicas es un aparato de IRM;
- Mamografía → el dispositivo de formación de imágenes médicas es un mamógrafo de rayos X;

- Etc.

La adquisición de dicha imagen puede implicar la inyección de un agente de contraste tal como gadolinio (GBCA) para IRM o agentes de contraste de rayos X apropiados.

Obsérvese que las “imágenes” pueden ser objetos 2D (con dos dimensiones espaciales), pero también posiblemente objetos 3D (con tres dimensiones espaciales), es decir, volúmenes constituidos de pilas de imágenes bidimensionales como “rebanadas” según una tercera dimensión espacial - en otras palabras, tenemos 2+1 dimensiones espaciales).

Además, podemos tener una pluralidad de imágenes de precontraste y/o imágenes candidatas de contraste.

En una realización preferida de IRM, se tienen al menos una imagen de precontraste ponderada en T1 y una imagen de contraste ponderada en T1, en particular imágenes GRE (eco de gradiente) de precontraste/contraste ponderadas en T1, pero se puede tener además:

- una imagen de precontraste ponderada en T2-flair (en particular TSE (turbo espín eco)); y/o
- un mapa de precontraste de ADC (coeficiente de difusión aparente) (convertido automáticamente a partir de EPI (ecoplanar) DWI (ponderada en difusión)).

Obsérvese que también puede disponerse de una imagen de contraste ponderada en T1 TSE (turboespín eco), pero no se usará en este caso.

Los métodos antes mencionados se implementan dentro de una arquitectura como la ilustrada en la **Figura 1**, por medio de un primer y/o segundo servidor 1a, 1b. El primer servidor 1a es el servidor de entrenamiento (que implementa un método de entrenamiento de la CNN) y el segundo servidor 1b es un servidor de procesamiento (que implementa el método de procesamiento). Es totalmente posible que estos dos servidores se fusionen.

Cada uno de estos servidores 1a, 1b es normalmente un equipo informático remoto conectado a una red extendida 2, tal como internet, para el intercambio de datos. Cada uno de ellos comprende medios de procesamiento de datos 11a, 11b de tipo procesador (en particular, el procesador de datos 11a del primer servidor 1a tiene una gran potencia de cálculo, ya que el aprendizaje es largo y complejo en comparación con el uso ordinario de los modelos entrenados), y opcionalmente medios de almacenamiento 12a, 12b, tales como una memoria de ordenador, por ejemplo, un disco duro. El segundo servidor 1b puede estar

conectado a uno o más dispositivos de formación de imágenes médicas 10 como equipo cliente, para proporcionar imágenes a procesar y recibir parámetros de vuelta.

Obsérvese que se supone que el dispositivo de formación de imágenes 10 comprende un inyector para realizar la inyección del agente de contraste, aplicando dicho inyector los parámetros de inyección.

La memoria 12a del primer servidor 1a almacena una base de datos de entrenamiento, es decir, un conjunto de imágenes denominadas imágenes de entrenamiento (en contraposición a las denominadas imágenes de entrada que precisamente se desea procesar). Cada imagen de la base de datos puede ser de precontraste o de contraste, y las imágenes de contraste pueden etiquetarse en función de la dosis de agente de contraste inyectada. Obsérvese que las imágenes correspondientes a una misma inyección (es decir, que forman una secuencia temporal) se asocian en secuencias.

CNN de minimización de dosis

En la siguiente descripción, se hará referencia a una “dosis baja” y a una “dosis habitual” de agente de contraste inyectado. Ambas están predeterminadas. Por comodidad, una imagen de contraste que representa una parte del cuerpo después de una inyección de la dosis habitual del agente de contraste será referida como “imagen de contraste de dosis habitual”, y una imagen de contraste que representa una parte del cuerpo después de una inyección de la dosis baja del agente de contraste será referida como “imagen de contraste de dosis baja”.

La dosis habitual, o “dosis completa”, es la dosis recomendada que se usa generalmente para un examen médico. Aunque los efectos secundarios son posibles, dicha dosis no es particularmente peligrosa para el paciente, y permite un nivel de calidad de imagen suficiente para fines de análisis/diagnóstico. En el caso de la GBCA, la dosis habitual es de 0.1 mmol/kg.

La dosis baja, o “dosis reducida”, es una dosis inferior a la dosis habitual, y que causa menos efectos en la salud del paciente. Dicha dosis baja puede ser cualquier fracción de la dosis habitual y es preferiblemente entre 1/10 y 1/2 (entre 10 y 50 %) de la dosis habitual, preferiblemente entre 1/5 y 1/3 (entre 20 y 33 %) de la dosis habitual, preferiblemente aproximadamente 1/4 (25 %). En el caso de GBCA, la dosis baja es, por ejemplo, 0.025 mmol/kg (25 % de 0.1 mmol/kg).

Obviamente, la imagen de contraste a dosis baja presenta menos contraste que la imagen de contraste a dosis habitual. Y amplificar ingenuamente el realce de contraste de un $x\%$ de CE-IRM de dosis baja por un factor de $100/x$ resulta en una calidad de imagen pobre con ruido generalizado y estructuras ambiguas.

La presente CNN se entrena para reconstruir, a partir de al menos una imagen de entrada de precontraste y una imagen de entrada de contraste de dosis baja que representan respectivamente una parte del cuerpo antes y después de una inyección de dicha dosis baja de agente de contraste, una imagen de contraste de dosis habitual que representa dicha parte del cuerpo después de una inyección de dicha dosis habitual de agente de contraste.

Dichas CNN son conocidas por el experto, véanse, por ejemplo, la solicitud WO2019074938 o el documento Ammari S, Bône A, Balleyguier C, et al. Can Deep Learning Replace Gadolinium in Neuro-Oncology? Invest Radiol. 2021; Publish Ah(00):1-9. doi: 10.1097/rli.000000000000008, y permiten reducir la dosis de agente de contraste: al aplicar la CNN a una imagen candidata de precontraste y a una imagen candidata de contraste de dosis baja que representan una parte del cuerpo antes y después de una inyección de sólo la dosis baja de agente de contraste, se puede generar una imagen de contraste sintética que simula dicha parte del cuerpo después de una inyección de la dosis habitual.

En otras palabras, esta CNN es capaz de predecir la imagen de contraste de dosis habitual a partir de la imagen de precontraste y de la imagen de contraste de dosis baja, y por lo que puede denominarse “CNN de minimización de dosis”.

De hecho, se puede seguir teniendo la calidad de la imagen de contraste de dosis habitual mientras sólo se inyecta la dosis baja de agente de contraste.

A continuación se describirán posibles formas de realización de esta CNN y sus métodos de entrenamiento.

Preferentemente, dicha CNN usa como entradas al menos las imágenes de precontraste/contraste ponderadas en T1 (y posiblemente también la imagen de precontraste ponderada en T2-flair y/o el mapa de precontraste de ADC) y produce otra imagen de contraste ponderada en T1.

El presente método propone un uso inteligente de esta CNN de minimización de dosis para aumentar virtualmente la dosis, al aplicar la CNN de reducción de dosis sobre la imagen candidata de contraste que representa una parte del cuerpo después de una inyección de

una “primera dosis” superior a la dosis baja de agente de contraste, posiblemente hasta la dosis habitual (pero no superior). Esto conduce a la generación de una imagen de contraste sintética que se cree representativa de lo que sería dicha parte del cuerpo tras una inyección de una “segunda dosis” de agente de contraste que es superior a la dosis habitual, es decir, una “superdosis”.

En consecuencia, se obtiene una imagen de muy alta calidad, como si se inyectara una dosis elevada (que podría ser peligrosa para el paciente), pero sin inyectar en realidad más que la dosis habitual de agente de contraste.

En otras palabras, la CNN no se usa para la tarea para la que realmente ha sido entrenada, lo cual es muy inusual, véanse las dos canalizaciones a modo de ejemplo de la figura 2 en donde la imagen de contraste de dosis habitual se usa directamente como imagen de contraste de primera dosis.

En esta figura, la imagen de precontraste ponderada en T1 se denomina “T1”, la imagen de precontraste ponderada en T2-flair se denomina “T2”, el mapa de precontraste de ADC se denomina “ADC”, la imagen de contraste ponderada en T1 de dosis baja se denomina “T1c-“, la imagen de contraste ponderada en T1 de dosis habitual se denomina “T1c”, la imagen de contraste ponderada en T1 de “superdosis” se denomina “T1c+”.

Partiendo de la hipótesis de que la CNN aprendió principalmente a amplificar la diferencia de contraste entre sus entradas de precontraste y poscontraste, cabía esperar que la sustitución de imágenes T1c- por imágenes T1c en la inferencia sintetizara imágenes T1c+ de dosis cuádruple aproximada.

Por ejemplo, suponiendo que la dosis baja (D_c^-) es $x\%$ de la dosis habitual (D_c): $D_c^- = x/100 * D_c$

Dado que se usa la misma CNN, la primera y la segunda dosis (D_1 , D_2) pueden escribirse con respecto a la misma ecuación: $D_1 = x/100 * D_2$.

$$\text{Y así } D_c^- < D_1 \leq D_c$$

$$x/100 * D_c < x/100 * D_2 \leq D_c$$

$$D_c < D_2 \leq 100/x * D_c, \text{ con } D_c^+ = 100/x * D_c$$

Por consiguiente, en la realización en donde la dosis baja es el 25% de la dosis habitual, la segunda dosis puede alcanzar el 400 % de la dosis habitual si la dosis habitual se usa como primera dosis (véase la figura 2). En otras palabras, si la dosis habitual es de 0.1

mmol.kg, la segunda dosis puede alcanzar 0.4 mmol/kg (inyección suplementaria virtual de 0.3 mmol/kg), de ahí el nombre de “superdosis”.

Obsérvese que dicha primera dosis es ventajosamente al menos el 80 % de la dosis habitual, preferiblemente aproximadamente el 100% de la dosis habitual, lo que significa que la segunda dosis es normalmente al menos el 160 % de la dosis habitual (80 %/50 %), preferiblemente al menos el 240 % (80 %/33 %), y preferiblemente entre el 320 % y el 400 % de la dosis habitual (80 %/25 % y 100 %/25 %).

UNet tridimensional

La figura 3 representa una arquitectura a modo de ejemplo de CNN de minimización de dosis, del tipo U-Net. Obsérvese que en el caso de imágenes tridimensionales de precontraste/contraste, la CNN es a su vez una red tridimensional, como en el ejemplo de la figura 2, que maneja mapas de características cuatridimensionales.

U-Net es una red neuronal del tipo codificador-decodificador: comprende una rama codificadora (o “ruta de contracción”) que mapea la entrada (al menos una imagen de precontraste y la imagen de dosis baja/imagen de primera dosis) en una representación de alto nivel y, a continuación, una rama decodificadora (o “ruta de expansión”) que genera la imagen de salida (la imagen de dosis habitual/imagen de segunda dosis) a partir de la representación de alto nivel.

U-Net comprende además conexiones de salto (o “laterales”) entre la rama codificadora y la rama decodificadora.

La rama codificadora actúa como esqueleto, y puede considerarse como una red convencional de extracción de características que puede ser de muchos tipos, y en particular una CNN convencional, preferiblemente una red neuronal totalmente convolucional (sucesión directa de bloques de capas de convolución y capas no lineales tales como ReLU (unidad lineal rectificadora), que en particular alterna bloques de convolución residuales y estriados para reducir la muestra). La rama codificadora extrae de la imagen de entrada una pluralidad de mapas de características iniciales representativos de la imagen de entrada a diferentes escalas. Más concretamente, el esqueleto consiste en una pluralidad de bloques de convolución sucesivos, de tal manera que el primer bloque produce un primer mapa de características iniciales a partir de la entrada, a continuación el segundo bloque produce un segundo mapa de características iniciales a partir del primer mapa de características iniciales,

etc.

Para las redes neuronales convolucionales se entiende convencionalmente que la escala es menor con cada mapa sucesivo (en otras palabras, la resolución disminuye, el mapa de características se hace “más pequeño” y, por tanto, menos detallado), pero de mayor profundidad semántica, ya que se han capturado estructuras de la imagen de nivel cada vez más alto. En concreto, los mapas de características iniciales tienen un número creciente de canales a medida que disminuye su tamaño espacial.

En la práctica, se coloca una capa de agrupación entre dos bloques para disminuir la escala en un factor de 2 (normalmente convolución $2 \times 2 \times 2$ con paso 2 para el muestreo descendente en el caso de imágenes 3D), y de un bloque a otro se aumenta (y preferiblemente se duplica) el número de filtros de las capas de convolución usadas (generalmente convoluciones $3 \times 3 \times 3$).

En la U-Net habitual de 5 niveles hay, por ejemplo, números de canales sucesivos de 32, 64, 128, 256 y 512, y tamaños espaciales de mapa sucesivos (para una imagen de entrada de $160 \times 192 \times 160$) de $160 \times 192 \times 160$, $80 \times 96 \times 80$, $40 \times 48 \times 40$, $20 \times 24 \times 20$, $10 \times 12 \times 10$. Se observa que la entrada ya tiene 4 canales - 3 imágenes de precontraste (T1, T2-flair ADC) y 1 imagen de contraste (T1c- en el entrenamiento y T1c en funcionamiento), mientras que la salida como un solo canal (T1c en el entrenamiento y T1c+ en funcionamiento).

Se dice que los mapas de características obtenidos por la rama codificadora son iniciales debido a que serán reprocesados por la rama decodificadora. De hecho, como se ha explicado, los mapas de “bajo nivel” tienen una mayor resolución espacial, pero una menor profundidad semántica. La rama decodificadora pretende aumentar su profundidad semántica incorporando la información de los mapas de “alto nivel”.

Así, dicha rama decodificadora de la CNN tiene la arquitectura simétrica de la rama codificadora, ya que genera, a partir de los mapas de características iniciales, una pluralidad de mapas de características enriquecidos que vuelven a ser representativos de la imagen de entrada a diferentes escalas, pero incorporan la información de los mapas de características iniciales de menor o igual escala a la vez que reducen el número de canales.

En otras palabras, la rama decodificadora también consiste en una pluralidad de bloques de convolución sucesivos, pero en orden inverso, de manera que el primer bloque produce el primer mapa de características enriquecido (a partir del cual puede generarse

directamente la imagen de salida) a partir del segundo mapa de características enriquecido y del primer mapa de características inicial, después de que el segundo bloque produzca el segundo mapa de características enriquecido a partir del tercer mapa de características enriquecido y del segundo mapa de características inicial, etc. La rama decodificadora es también preferiblemente una CNN totalmente convolucional (sucesión directa de bloques de capas de convolución y capas no lineales, tales como ReLU, que en particular alterna bloques de convolución residual y estriada para aumentar la muestra).

En más detalles, cada i -ésimo mapa enriquecido tiene la escala del i -ésimo mapa inicial correspondiente (es decir, sensiblemente el mismo tamaño espacial), pero incorpora la información de todos los j -ésimo mapas, para cada $j \geq i$. En la práctica, cada i -ésimo mapa enriquecido D_i se genera en función del correspondiente i -ésimo mapa inicial E_i y/o del siguiente ($i+1$ -ésimo) mapa enriquecido, de ahí la naturaleza “contractiva y expansiva” de las ramas (es decir, en forma de “U”): los mapas iniciales se obtienen en orden ascendente y, a continuación, los mapas enriquecidos se obtienen en orden descendente.

De hecho, el nivel semántico máximo se obtiene en el mapa de “escala más pequeña”, y a partir de ahí cada mapa se enriquece de nuevo en sentido descendente con la información de los mapas ya enriquecidos. Las conexiones de salto entre la rama codificadora y la rama decodificadora proporcionan a la rama decodificadora los distintos mapas iniciales.

Normalmente, la generación de un mapa enriquecido basado en el mapa inicial correspondiente y el mapa enriquecido de escala más pequeña comprende el reescalado del mapa enriquecido, normalmente duplicando la escala (si ha habido reducción a la mitad de la escala en la rama codificadora), es decir, muestreo ascendente del mapa de características enriquecido mediante una convolución $2 \times 2 \times 2$ (“convolución ascendente”) que reduce a la mitad el número de canales de características, luego concatenación con el mapa inicial correspondiente E_i (recortado en caso necesario, ambos mapas tienen ahora sensiblemente la misma escala) para duplicar de nuevo el número de canales, y de un bloque a otro el número de filtros de las capas de convolución usadas (generalmente convoluciones $3 \times 3 \times 3$) se reduce de nuevo (y preferiblemente se reduce aún más a la mitad).

Los bloques de convolución de muestreo descendente y muestreo ascendente se basan en núcleos $2 \times 2 \times 2$, todos los demás núcleos son $3 \times 3 \times 3$, a excepción de la convolución

final $1 \times 1 \times 1$. Todas las funciones de activación excepto la última sigmoide son 0.2-LeakyReLU.

Obsérvese que la presente invención no está limitada a la arquitectura específica de U-Net: podría ser adecuada cualquier CNN que sea entrenada para reconstruir, a partir de al menos una imagen de precontraste de entrenamiento y una imagen de contraste de entrenamiento que representen respectivamente una parte del cuerpo antes y después de una inyección de dicha dosis baja de agente de contraste, una imagen de contraste de entrenamiento que represente dicha parte del cuerpo después de una inyección de dicha dosis habitual de agente de contraste.

El entrenamiento se describirá más adelante.

Método para la obtención de imágenes médicas

Como se representa en la figura 4, el presente método comienza con una etapa (a) de obtención de dicha imagen de precontraste e imágenes de contraste a procesar (denominadas imágenes candidatas de precontraste y contraste), en particular imágenes de precontraste/contraste ponderadas en T1, preferiblemente a partir de un dispositivo de formación de imágenes médicas 10 conectado al segundo servidor 1b.

La imagen candidata de precontraste y la imagen candidata de contraste representan respectivamente una parte del cuerpo antes y después de una inyección de una primera dosis de agente de contraste, en donde dicha primera dosis es superior a una dosis baja predeterminada, siendo la dosis baja predeterminada inferior a una dosis habitual predeterminada.

En la realización preferida de IRM, la etapa (a) comprende la obtención de tres imágenes candidatas de precontraste (la imagen candidata de precontraste ponderada en T1, una imagen candidata de precontraste ponderada en T2-flair y un mapa candidato de precontraste de ADC).

Obsérvese que esta etapa (a) puede ser implementada por el procesador de datos 11b del segundo servidor 1b y/o por el dispositivo de formación de imágenes médicas 10.

En una etapa principal (b), implementada por el procesador de datos 11b del segundo servidor 1b, la CNN se aplica a la imagen de precontraste candidata y a la imagen de contraste candidata, para generar una imagen de contraste sintética que se cree que simula dicha parte del cuerpo después de una inyección de una segunda dosis de agente de

contraste que es mayor que la dosis habitual.

En la realización de IRM preferida, la CNN se aplica a las tres imágenes candidatas de precontraste y a la imagen candidata de contraste.

Método de entrenamiento

El método comprende ventajosamente una etapa previa (a0) de entrenamiento de la CNN, implementada por el procesador de datos 11a del primer servidor 1a.

Por entrenamiento se entiende la determinación de los valores óptimos de parámetros y pesos de la CNN. Obsérvese que, alternativamente, la CNN puede tomarse directamente “de las disponibles” con valores preestablecidos de parámetros y pesos.

Dicho método de entrenamiento puede realizarse según el estado de la técnica, y puede usarse cualquier protocolo de entrenamiento adecuado conocido por un experto.

Debe entenderse que la CNN se entrena para realizar su tarea original de “minimización de dosis”, es decir, para reconstruir, a partir de al menos una imagen de precontraste de entrenamiento y una imagen de contraste de entrenamiento que representan respectivamente una parte del cuerpo antes y después de una inyección de dicha dosis baja de agente de contraste, una imagen de contraste de entrenamiento que representa dicha parte del cuerpo después de una inyección de dicha dosis habitual de agente de contraste.

De hecho, es imposible entrenar directamente una CNN para generar, a partir de al menos una imagen de precontraste de entrenamiento y una imagen de contraste de entrenamiento que representen respectivamente una parte del cuerpo antes y después de una inyección de una primera dosis de agente de contraste superior a la dosis inferior, una imagen de contraste sintética que simule dicha parte del cuerpo después de una inyección de una segunda dosis de agente de contraste superior a la dosis habitual: dicho entrenamiento requeriría disponer de imágenes de entrenamiento después de una inyección de la segunda dosis, con lo que la cantidad total de dosis inyectada sería superior a la cantidad de dosis sugerida por las directrices clínicas recientes.

Por el contrario, para aprender la tarea de “minimización de dosis”, la base de entrenamiento sólo tiene que comprender una pluralidad de secuencias de al menos una imagen de precontraste de entrenamiento (como se ha explicado podría haber posiblemente tres imágenes de precontraste), una imagen de contraste de entrenamiento “inicial” (imagen de contraste de dosis baja) y una imagen de contraste de entrenamiento “final” (imagen de

contraste de dosis habitual) que representan respectivamente una parte del cuerpo antes, después de una inyección de dicha dosis baja de agente de contraste, y después de una inyección de dicha dosis habitual de agente de contraste.

La CNN se entrena para predecir la imagen de contraste de entrenamiento final (como verdad absoluta) a partir de la(s) imagen(es) de precontraste de entrenamiento y la imagen de contraste inicial de la misma secuencia.

En más detalle, para generar un ejemplo de entrenamiento, realizamos el protocolo normal para adquirir una imagen de contraste que representa una parte del cuerpo después de una inyección de dicha dosis habitual de agente de contraste, pero la administración de la dosis habitual simplemente se divide en dos inyecciones sucesivas (1) la dosis baja y (2) la diferencia entre la dosis baja y la dosis habitual, por ejemplo 0.025 mmol/kg y $0.1 - 0.025 = 0.075 \text{ mmol/kg}$ de GBCA como se ha explicado, y se adquiere una imagen de contraste adicional entre ambas como imagen de contraste “inicial”, esta imagen en realidad representa la parte del cuerpo después de una inyección de dicha dosis baja de agente de contraste. Obsérvese que recientemente se ha incluido protocolos de IRM de dos inyecciones en las directrices de consenso para la obtención de imágenes de gliomas, desempeñando la primera inyección el papel de bolo de precarga para una posible secuencia de perfusión posterior. Por consiguiente, es inocua.

En resumen, para un paciente

- se adquiere(n) la(s) imagen(es) de precontraste;
- se inyecta la dosis baja y se adquiere una primera imagen de contraste;
- se inyecta la diferencia entre la dosis habitual y la dosis baja (para que el paciente reciba en total la dosis habitual), y se adquiere una segunda imagen de contraste;
- la secuencia de la(s) imagen(imágenes) de precontraste, las dos imágenes de contraste sucesivas se etiquetan como ejemplo de entrenamiento.

Pruebas

En la Figura 5 se muestran las imágenes T2-flair, T1c, T1c+ y tse-T1c de cuatro exámenes de RMN de ejemplo de la muestra de prueba, con metástasis cerebrales o glioma. Podemos ver que las imágenes T1c+ (las imágenes de contraste con superdosis sintéticas) presentan una calidad excelente y son más fáciles de leer.

La Tabla 1 recoge las puntuaciones medias de calidad de imagen (IQ) de las

imágenes T1c, T1c+ y tse-T1c de 79 exámenes de una muestra de prueba. Las calificaciones se expresan en una escala de Likert de 4 puntos que va de 1 (deficiente) a 4 (excelente). Las desviaciones estándar se indican entre paréntesis. Las diferencias entre los lectores y las imágenes de IRM poscontraste se comparan usando pruebas de la t bilaterales y se indican los valores de p. Los mejores parámetros se resaltan con negrita cuando se alcanza el umbral de significación del 5%).

Esta tabla muestra que las imágenes T1c+ sintetizadas fueron las preferidas por dos lectores (neurorradiólogos con 10 y 11 años de experiencia respectivamente) por su calidad de imagen general, cuando no se encontraron diferencias de calidad entre las imágenes T1c y tse-T1c. Por término medio entre los lectores, las imágenes T1c y tse-T1c recibieron una calificación de 2.7/4 (de media a buena), mientras que las imágenes T1c+ recibieron una calificación de 3.4/4 (de buena a excelente).

	IQ $\pm$ est. (grados en una escala de 4 puntos)			T1c frente a T1c+	T1c frente a tse-T1C	T1c+ frente a tse-T1C
	T1c	T1c+	tse-T1c			
Lector promedio	2.7 ($\pm$ 0.5)	3.4 ($\pm$ 0.3)	2.7 ($\pm$ 0.6)	$P<0.001$ *	$P=0.62$	$P<0.001^*$
Lector 1	2.8 ($\pm$ 0.5)	2.9 ($\pm$ 0.4)	2.7 ($\pm$ 0.6)	$P=0.02^*$	$P=0.36$	$P=0.007^*$
Lector 2	2.6 ($\pm$ 0.7)	3.9 ($\pm$ 0.4)	2.7 ($\pm$ 0.8)	$P<0.001$ *	$P=0.14$	$P<0.001^*$
R1 frente a R2	$P=0.007^*$	$P<0.001^*$	$P=0.89$			

Tabla 1

En la Tabla 2 se presentan los parámetros promedio de la relación contraste-ruído (CNR), la relación lesión-cerebro (LBR) y el porcentaje de mejora del contraste (CEP) de las imágenes T1c, T1c+ y tse-T1c de 52 exámenes de la muestra de prueba con al menos una lesión de referencia. Las desviaciones estándar se indican entre paréntesis. Las diferencias

entre los lectores y las imágenes de IRM poscontraste se comparan usando pruebas de la t bilaterales y se indican los valores de p. Se resaltan en negrita los mejores parámetros cuando se alcanza el umbral de significación del 5 %.

Con una CNR media de 44.5, LBR de 1.66 y CEP de 112.4 %, las imágenes T1c+ sintetizadas superan tanto a las imágenes T1c como a las tse-T1c en todos los parámetros considerados.

	T1c	T1c+	tse-T1c	T1c frente a T1c+	T1c frente a tse-T1C	T1c+ frente a tse-T1C
CNR	9.1 ($\pm$ 5.8)	44.5 ($\pm$ 25.7)	16.8 ($\pm$ 9.7)	$P<0.001^*$	$P<0.001^*$	$P<0.001^*$
LBR	1.31 ($\pm$ 0.18)	1.66 ($\pm$ 0.27)	1.44 ($\pm$ 0.24)	$P<0.001^*$	$P<0.001^*$	$P<0.001^*$
CEP	85.6 ($\pm$ 35.5)	112.4 ($\pm$ 36.5)	92.2 ($\pm$ 38.1)	$P<0.001^*$	$P=0.011$	$P<0.001^*$

Tabla 2

Finalmente, se comparó el rendimiento promedio de detección de lesiones alcanzado cuando se leen las imágenes T1c, y cuando se lee también conjuntamente su homóloga T1c+ posprocesada. En ambos escenarios de lectura, los lectores disponían también de la imagen T2-flair precontraste. Una lectura con acceso a imágenes T2-flair, T1c y tse-T1c definió 187 lesiones de referencia con una mediana de longitud del eje largo de 9.2 mm (intervalo intercuartílico de 10.7 mm). Se consideraron tres configuraciones de evaluación anidadas, dependiendo del tamaño mínimo de la lesión incluida: 10 mm, 5 mm o 0 mm, considerándose en este caso todas las lesiones.

El acceso a imágenes T1c+ aumentó la sensibilidad de detección de lesiones (SE) para ambos lectores en todas las configuraciones de evaluación. De media para todos los lectores, la SE global fue del 75 % cuando se disponía de imágenes T1c+ (59 % sólo con imágenes T1c, $P<0.001^*$), del 85% para lesiones mayores de 5 mm (70 % con T1c, $P<0.001^*$), y del 96 % para lesiones mayores de 10 mm (88 % con T1c, $P=0.008^*$). No se encontraron diferencias en términos de FDR (tasas de falsa detección), que se mantuvieron por debajo de 0.19/examen en todos los lectores, escenarios de lectura y configuraciones de evaluación. No se encontraron diferencias entre los lectores ni para SE ni para FDR.

El VPP se mantuvo por encima del 90 % en todas las configuraciones. De media entre los lectores, las diferencias cuantitativas del VPP fueron del 3 % como máximo entre los dos escenarios de lectura. La F1 fue superior al 90 % en todos los lectores y lecturas para las

lesiones mayores de 10 mm, superior al 80 % para las lesiones mayores de 5 mm y superior al 70 % cuando se incluyeron todas las lesiones. En promedio entre los lectores, la F1 fue superior cuando los lectores disponían de T1c+ por un margen que varió entre el 4 % y el 11 % según el intervalo considerado de tamaños de lesión.

Producto del programa informático

En un segundo y un tercer aspecto, la invención proporciona un producto de programa informático que comprende instrucciones de código para ejecutar un método (en particular en el procesador de datos 11a, 11b del primer y/o segundo servidor 1a, 1b) según el primer segundo aspecto de la invención para la obtención de imágenes médicas, y medios de almacenamiento legibles por equipos informáticos (memoria del primer o segundo servidor 1a, 1b) provistos de este producto de programa informático.

Se hace constar que con relación a esta fecha, el mejor método conocido por la solicitante para llevar a la práctica la citada invención, es el que resulta claro de la presente descripción de la invención.