(19) LE GOUVERNEMENT
DU GRAND-DUCHÉ DE LUXEMBOURG
Ministère de l'Économie

(11) N° de publication : **LU506063**

(12) **BREVET D'INVENTION** **B1**

(21) N° de dépôt: LU506063

(22) Date de dépôt: 09/01/2024

(51) Int. Cl.:
G06V 20/69, G06T 7/00, G16H 10/40

(30) Priorité:

(43) Date de mise à disposition du public: 09/07/2024

(47) Date de délivrance: 09/07/2024

(73) Titulaire(s):
JIAXING VOCATIONAL & TECHNICAL COLLEGE – Jiaxing City, Zhejiang, (Chine)

(72) Inventeur(s):
SHEN Xudong – Chine

(74) Mandataire(s):
IP SHIELD – 1616 Luxembourg (Luxembourg)

(54) **EIN VERFAHREN ZUR AUTOMATISCHEN KLASSIFIZIERUNG VON ZELLEN IN GEBÄRMUTTERHALSFLÜSSIGKEIT.**

(57) Die vorliegende Erfindung bezieht sich auf das Gebiet der Zelldetektionstechnologie, insbesondere auf ein Verfahren zur automatischen Klassifizierung und Detektion von Zellen auf der Basis von Gebärmutterhalsflüssigkeit, das die folgenden Schritte umfasst: Aufbau eines CPCA-Mechanismus für räumliche Aufmerksamkeit unter Verwendung eines tiefen Faltungsmoduls und Erzeugung einer dynamisch verteilten räumlichen Aufmerksamkeitskarte auf jedem Kanal; Um die Zielvielfalt zu erhöhen und das Problem unzureichender Datensätze und unausgewogener Proben für die Erkennung kleiner Ziele zu lösen, wird zusätzlich zur Mosaik-Datenverbesserung, affinen Transformation und perspektivischen Transformation das Mixup- Modul hinzugefügt, um die Datenverbesserung durchzuführen, d. h. zwei Bilder werden zufällig aus jedem Batch ausgewählt und in einem bestimmten Verhältnis gemischt, um ein neues Bild zu erzeugen; Der vorteilhafte Effekt ist: die automatische Klassifizierung und Erkennung von Zellen auf der Basis von Gebärmutterhalsflüssigkeit, die durch die vorliegende Erfindung vorgeschlagen wird, integriert das CPCA-Modul auf der Grundlage des ursprünglichen Yolov5s- Netzwerkmodells, das das Netzwerkmodell zur Erkennung pathologischer Merkmale abnormaler Gebärmutterhalszellen verbessert, indem es die Kanal- und Positionsinformationen erfasst und die wichtigen Merkmale effektiv hervorhebt, und es hat einen geringeren Berechnungsaufwand.

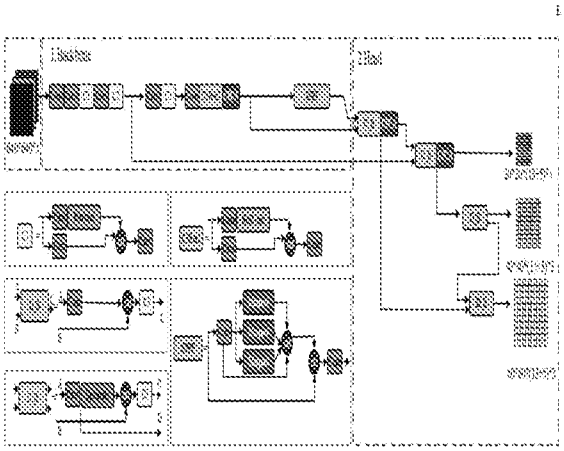

Bild 1

LU506063

Ein Verfahren zur automatischen Klassifizierung von Zellen in Gebärmutterhalsflüssigkeit

Technischer Bereich

Die vorliegende Erfindung bezieht sich auf das Gebiet der Zelldetektionstechnologie, insbesondere auf ein Verfahren zur automatischen Klassifizierung von Zellen in Gebärmutterhalsflüssigkeit.

Technologie im Hintergrund

Gebärmutterhalskrebs ist die zweithäufigste Krebserkrankung bei Frauen weltweit und ein häufiger bösartiger Tumor bei Frauen. Einschlägige Daten zeigen, dass die Inzidenz von Gebärmutterhalskrebs in China sechsmal höher ist als in den Industrieländern. Im Jahr 2020 wird es in China etwa 110.000 neue Fälle von Gebärmutterhalskrebs und etwa 59.000 Todesfälle geben, was etwa 18 Prozent der weltweiten Todesfälle ausmacht.

Gebärmutterhalskrebs ist nach dem Stand der Technik vermeidbar, und eine frühzeitige Diagnose ist entscheidend für die Verbesserung der Überlebensrate bei Gebärmutterhalskrebs. Zytologische Screeningverfahren sind seit mehr als einem halben Jahrhundert eine wichtige Methode der Gebärmutterhalskrebsdiagnose und -prävention, die die Diagnose von präkanzerösen und krebsartigen Läsionen erheblich verbessert hat. Die Zytologie des Gebärmutterhalses wird in der Regel von einem Zytopathologen anhand eines herkömmlichen Pap-Abstrichs oder eines flüssigkeitsbasierten Zytologie-Objektträgers diagnostiziert, der zunächst abnorme Läsionszellen aus Tausenden von Zellen identifiziert und dann den Abstrich sortiert, um einen zytologischen Bericht zu erstellen. Die manuelle Untersuchung zervikaler zytologischer Abstriche ist jedoch gleichbedeutend mit der Suche nach der Nadel im Heuhaufen, die sehr mühsam und kompliziert und anfällig für subjektive Fehler ist; um die Effektivität und Effizienz des zervikalen zytologischen Screenings weiter zu verbessern, besteht daher ein extrem hoher Bedarf an der Entwicklung einer automatisierten Analyse der zervikalen Zytologie. Es gibt viele computergestützte Analysemethoden zur Erkennung abnormaler oder kranker Zellen auf Bildern von Gebärmutterhalszellen. Die Erkennung von Zellen aus Gebärmutterhalsflüssigkeiten in der traditionellen Medizin weist viele Mängel auf, und die Methoden der Zellerkennung werden schrittweise verbessert. Mit der kontinuierlichen Entwicklung der Deep-Learning-Technologie sind viele klassische und fortschrittliche Algorithmen entstanden, wie Fast R-CNN, ResNet, YOLOV3 und YOLOV5. Diese Deep-Learning-Algorithmen können die Aufgaben der Klassifizierung und Erkennung erfüllen, indem sie die Merkmale der Bilddaten extrahieren und die komplexen Merkmalsbeziehungen zwischen den Bildern lernen.

Für den Nachweis abnormaler Gebärmutterhalszellen können normale und abnormale Zellen jedoch ein sehr ähnliches Aussehen haben, so dass eine solche lokalisierte Argumentation oft nicht ausreicht. In der klinischen Medizin vergleicht der Zytopathologe, um festzustellen, ob eine Gebärmutterhalszelle normal ist oder nicht, in der Regel die sie umgebenden Referenzzellen, um ihre Kategorie zu bestimmen. Die Kategorisierung von Gebärmutterhalszellen kann sich nicht allein auf Merkmale stützen, die aus der Zielregion oder der Region von Interesse der Zellen extrahiert werden. Bestehenden Methoden fehlt es oft an Merkmalen, die zwischen den Zellen interagieren, was zu einer schlechten Klassifizierungsleistung führt.

Inhalt der Erfindung

Ziel der vorliegenden Erfindung ist es, ein Verfahren zur automatischen Klassifizierung von Zellen in Gebärmutterhalsflüssigkeit bereitzustellen, um die oben genannten Probleme des Standes der Technik zu lösen.

Um das obige Ziel zu erreichen, bietet die vorliegende Erfindung die folgende technische Lösung: ein Verfahren zur automatischen Klassifizierung von Zellen in Gebärmutterhalsflüssigkeit, wobei das Verfahren die folgenden Schritte umfasst:

Konstruktion eines CPCA-Mechanismus für räumliche Aufmerksamkeit unter Verwendung eines tiefen Faltungsmoduls zur Erzeugung einer dynamisch verteilten räumlichen Aufmerksamkeitskarte auf jedem Kanal;

Um die Zielvielfalt zu erhöhen und das Problem unzureichender Datensätze und unausgewogener Stichproben bei der Erkennung kleiner Ziele zu lösen, wird zusätzlich zur Mosaikdatenverbesserung, affinen Transformation und perspektivischen Transformation das Mixup-Modul zur Datenverbesserung hinzugefügt, d. h. zwei Bilder werden zufällig aus jedem Batch ausgewählt und mit einem bestimmten Verhältnis gemischt, um ein neues Bild zu erzeugen;

Verbesserung der Verlustfunktion durch Ersetzen der Verlustfunktion „Generalised Intersection and Combination Ratio“ GIoU durch die Verlustfunktion „Complete Intersection and Combination Ratio“ CIoU, um die Leistung der Bounding-Box-Regression und der Lokalisierung zur Identifizierung von zervikalen pathologischen Zellfraktionen mit komplexer Umgebung zu verbessern.

Vorzugsweise wird das CPCA-Modul dem Backbone-Netz und dem Neck-Netz hinzugefügt und nach dem C3-Modul platziert. Während des Netzwerktrainings verbessert das CPCA-Modul das Erlernen von Informationen über kleine Zielmerkmale, indem es die Gewichte des Kanal- und des räumlichen Aufmerksamkeitsmoduls dynamisch anpasst.

Vorzugsweise führt der Mechanismus der Kanal-A-priori-Faltungsaufmerksamkeit CPCA nacheinander zwei Modulaufgaben der Kanalaufmerksamkeit CA und der räumlichen Aufmerksamkeit SA durchführt, wobei angenommen wird, dass das Eingangsbildmerkmal $F \in R^{C\times H\times W}$ ist. Eine eindimensionale Kanalaufmerksamkeitsmerkmalskarte $M_C \in R^{C\times 1\times 1}$ wird zunächst durch das Kanalaufmerksamkeitsmodul CA erhalten, und dann wird M_C Element für Element mit dem Eingangsmerkmal F multipliziert, um den Kanalaufmerksamkeitswert aus der räumlichen Dimension zu übertragen, um so ein besseres Bildmerkmal $F_C \in R^{C\times H\times W}$ mit Kanalaufmerksamkeit zu erhalten; Das räumliche Aufmerksamkeitsmodul SA erzeugt eine räumliche 3D-Aufmerksamkeitskarte $M_s \in R^{C\times H\times W}$ durch Datenverarbeitung der Bildmerkmale F_C, und gibt dann schließlich die verarbeiteten Bildmerkmale $\hat{F} \in R^{C\times H\times W}$ durch elementweise Multiplikation von M_s und F_C aus, und der gesamte Prozess wird durch die folgende Gleichung ausgedrückt:

$$F_C = CA(F) \otimes F$$

$$\hat{F} = SA(F_C) \otimes F_C$$

Vorzugsweise erzeugt das Kanalaufmerksamkeitsmodul eine Kanalaufmerksamkeitsmerkmalskarte, indem es die in den Bildmerkmalen vorhandenen Beziehungen zwischen den Kanälen untersucht; die räumliche Aufmerksamkeitskarte wird durch

Extraktion räumlicher Zuordnungsbeziehungen erzeugt.

Vorzugsweise verwendet das Mischmodul ein lineares Differenzverfahren anwendet, um zwei unterschiedliche Trainingsproben und Etikettenproben aus dem ursprünglichen Datensatz zum Mischen und Erzeugen einer neuen Probe auszuwählen, und dass zwei unterschiedliche Trainingsprobendaten (d_i, l_i) und (d_j, l_j) zufällig aus dem Trainingssatz ausgewählt werden.

Wobei d die Trainingsdaten, l die etikettierten Daten und $\mathrm{i} \neq j$ bezeichnet, dann werden die beiden Zufallsstichproben durch den folgenden Algorithmus berechnet, um eine neue Stichprobe zu erzeugen, der Algorithmus ist in der folgenden Gleichung angegeben:

$$\hat{\mathrm{d}} = \lambda \mathrm{d}_{\mathrm{i}} + (1-\lambda) d_j \ ;$$

$$\hat{l} = \lambda l_{\mathrm{i}} + (1-\lambda) l_j \ ;$$

Dabei ist $\lambda \in [0,1]$ ein Hyperparameter, der der $Beta \in (\alpha, \alpha)$ Verteilung gehorcht, d.h. der Intensitätskoeffizient, der die Interpolation steuert. Die Größe von α ist proportional zur Intensität der Interpolation, die die Mischtiefe der Bildmusterdaten beeinflusst, und wenn G=0 ist, werden die Daten nicht gemischt und Mixup funktioniert nicht.

Vorzugsweise wird CIoU als Regressionsverlustfunktion verwendet, wie in der folgenden Gleichung ausgedrückt:

$$L_{loc} = 1 - IoUI\ (B,\ B_{\mathrm{gt}}) + \frac{d^2}{c^2} + \alpha v \ ;$$

$$\alpha = \frac{v}{1 - IoU + v} \ ;$$

$$v = \frac{4}{\pi^2} (\arctan, \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h})^2 \ ;$$

Dabei steht w für die Breite des Vorhersagekastens, h für die Höhe des Vorhersagekastens, w^{gt} für die Breite des Zielkastens und h^{gt} für die Höhe der Zielbibliothek.

Verglichen mit dem Stand der Technik hat die vorliegende Erfindung folgende vorteilhafte Auswirkungen:

Das Verfahren zur automatischen Klassifizierung von Zellen in Gebärmutterhalsflüssigkeit, die durch die vorliegende Erfindung vorgeschlagen wird, verbessert das Netzwerkmodell zur Erkennung von pathologischen Merkmalen von abnormalen Gebärmutterhalszellen durch die Integration des CPCA-Moduls auf der Grundlage des ursprünglichen Yolov5s-Netzwerkmodells, hebt die wichtigen Merkmale durch die Erfassung der Kanal- und Positionsinformationen effektiv hervor und hat einen geringeren Rechenaufwand.

Beschreibung der beigefügten Zeichnungen

Bild 1 zeigt ein schematisches Diagramm der YOLOv5s-CPCA-Netzwerkstruktur der vorliegenden Erfindung;

Bild 2 zeigt ein schematisches Diagramm der Struktur des CPCA-

Aufmerksamkeitsmechanismus der vorliegenden Erfindung.

Detaillierte Beschreibung

Um den Zweck der vorliegenden Erfindung, die technischen Lösungen für eine klare und vollständige Beschreibung, und die Vorteile klarer zu verstehen, sind die folgenden Ausführungsformen der vorliegenden Erfindung im Detail in Verbindung mit den beigefügten Zeichnungen beschrieben. Es sollte verstanden werden, dass die spezifischen Ausführungsformen hierin beschrieben sind Teil der Ausführungsformen der vorliegenden Erfindung, nicht alle der Ausführungsformen, nur für die Erklärung der Ausführungsformen der vorliegenden Erfindung, und wird nicht verwendet, um die Ausführungsformen der vorliegenden Erfindung zu begrenzen, alle anderen Ausführungsformen durch den normalen Fachmann auf dem Gebiet erhalten, ohne dass kreative Arbeit unter der Prämisse des Schutzes des Anwendungsbereichs der vorliegenden Erfindung.

Ausführungsform 1

Unter Bezugnahme auf die Bilder 1 bis 2 bietet die vorliegende Erfindung eine technische Lösung: ein Verfahren zur automatischen Klassifizierung und Erkennung von Zellen auf der Basis von Gebärmutterhalsflüssigkeit, wobei das Verfahren die folgenden Schritte umfasst:

Aufbau eines CPCA-Mechanismus für räumliche Aufmerksamkeit unter Verwendung eines tiefen Faltungsmoduls zur Erzeugung einer dynamisch verteilten räumlichen Aufmerksamkeitskarte auf jedem Kanal;

Um die Zielvielfalt zu erhöhen und das Problem unzureichender Datensätze und unausgewogener Stichproben für die Erkennung kleiner Ziele zu lösen, wird zusätzlich zur Mosaik-Datenverbesserung, affinen Transformation und perspektivischen Transformation das Mixup-Modul hinzugefügt, um eine Datenverbesserung durchzuführen, d.h. zwei Bilder werden zufällig aus jedem Batch ausgewählt und in einem bestimmten Verhältnis gemischt, um ein neues Bild zu erzeugen;

Verbesserung der Verlustfunktion durch Ersetzen der Verlustfunktion „Generalised Intersection and Combination Ratio“ GIoU durch die Verlustfunktion „Complete Intersection and Combination Ratio“ CIoU, um die Leistung der Bounding-Box-Regression und der Lokalisierung zur Identifizierung von zervikalen pathologischen Zellfraktionen mit komplexer Umgebung zu verbessern.

Das YOLOv5s-Basismodell wurde so verbessert, dass es auch in der komplexen Umgebung von Gebärmutterhalszellen über starke Erkennungsfähigkeiten verfügt. Der Backbone-Netzteil des Modells wird verwendet, um Merkmale aus dem Bild auf verschiedenen feinkörnigen Ebenen zu extrahieren, und das ursprüngliche Backbone-Netz besteht aus Conv, C3, C3Res2 und Spatial Pyramid Pooling-Fast (SPPF). Die Größe des Eingangsbildes von Zellen auf der Basis von Gebärmutterhalsflüssigkeit im YOLOv5s-Modell ist so gewählt, dass die Details der Bildgenauigkeit vollständig erhalten bleiben. Die Größe des Eingangsbildes beträgt 1024x1024x3 und die Anzahl der Faltungskerne ist 32 in der ersten Focus-Struktur. Nach der Faltung beträgt die Größe der Merkmalskarte 512x512x32. Das Modul C3 wird verwendet, um die Merkmale der Gebärmutterhalszellen zu extrahieren, die im Backbone-Netzwerkteil und im Halsnetzwerkteil verwendet werden. Im Backbone-Netzwerkteil enthält das C3-Modul detaillierte Standortinformationen, aber weniger semantische Informationen; Im Teil des Halsnetzwerks wird das C3-Modul verwendet, um Merkmale zu extrahieren, die detaillierte semantische Informationen, aber weniger Positionsinformationen enthalten, und seine Verarbeitung von Merkmalsinformationen für kleine Ziele führt zu mehr fehlenden Positionsinformationen des Ziels,

wodurch leicht Merkmalsinformationen verloren gehen.SPPF wird verwendet, um Merkmalskarten verschiedener Größen als Ausgabe miteinander zu verbinden. Der Header-Vorhersageteil der Ausgabe sagt die verarbeiteten Gebärmutterhalszell-Bildmerkmale in drei verschiedenen Maßstäben voraus, erzeugt Bounding Boxes und sagt die Klasse des Ziels voraus. Um die oben genannten Probleme weiter zu verbessern und die Erkennungsgenauigkeit des ursprünglichen Modells für Gebärmutterhalszellen verschiedener Falltypen zu erhöhen, schlagen wir ein verbessertes YOLOv5s-Netzwerkmodell vor, wie in Abb. 1 dargestellt. Das CPCA-Modul wurde zu den Backbone- und Zervikalnetzwerken hinzugefügt und nach dem C3-Modul platziert. Das CPCA-Modul verbessert das Lernen von Informationen über kleine Zielobjekte, indem es die Gewichte des Kanal- und des räumlichen Aufmerksamkeitsmoduls während des Netzwerktrainings dynamisch anpasst.

Wie aus Bild 2 ersichtlich ist, führt der Mechanismus der Kanal-A-priori-Faltungsaufmerksamkeit CPCA nacheinander zwei Modulaufgaben der Kanalaufmerksamkeit CA und der räumlichen Aufmerksamkeit SA durchführt, wobei angenommen wird, dass das Eingangsbildmerkmal $F \in R^{C\times H\times W}$ ist. Eine eindimensionale Kanalaufmerksamkeitsmerkmalskarte $M_C \in R^{C\times 1\times 1}$ wird zunächst durch das Kanalaufmerksamkeitsmodul CA erhalten, und dann wird M_C Element für Element mit dem Eingangsmerkmal F multipliziert, um den Kanalaufmerksamkeitswert aus der räumlichen Dimension zu übertragen, um so ein besseres Bildmerkmal $F_C \in R^{C\times H\times W}$ mit Kanalaufmerksamkeit zu erhalten; Das räumliche Aufmerksamkeitsmodul SA erzeugt eine räumliche 3D-Aufmerksamkeitskarte $M_s \in R^{C\times H\times W}$ durch Datenverarbeitung der Bildmerkmale F_C, und gibt dann schließlich die verarbeiteten Bildmerkmale $\hat{F} \in R^{C\times H\times W}$ durch elementweise Multiplikation von M_s und F_C aus, und der gesamte Prozess wird durch die folgende Gleichung ausgedrückt:

$$F_C = CA(F) \otimes F$$

$$\hat{F} = SA(F_C) \otimes F_C$$

Kanalaufmerksamkeitsmodul (CA): Das Kanalaufmerksamkeitsmodul erzeugt eine Kanalaufmerksamkeitsmerkmalskarte, indem es die in den Bildmerkmalen vorhandenen Beziehungen zwischen den Kanälen untersucht. Zunächst werden die räumlichen Informationen aus dem Merkmalsbild extrahiert und die Operationen des durchschnittlichen Poolings und des maximalen Poolings werden unter Verwendung der im CBAM-Aufmerksamkeitsmechanismus vorgeschlagenen Methode angewandt; anschließend werden die aus dem durchschnittlichen Pooling und dem maximalen Pooling gewonnenen räumlichen Kontextinformationen an das gemeinsame mehrschichtige Perzeptron (MLP) gesendet. Dann werden die MLP-Ausgänge durch elementweise Summierung fusioniert, und schließlich erhält man die Merkmalskarten für die Kanalaufmerksamkeit. Um die Menge der Parameterdaten zu reduzieren, besteht die MLP-Struktur aus einer einzigen verborgenen Schicht, deren Größe auf $R^{c/r\times 1\times 1}$ festgelegt ist, wobei r

den Dämpfungskoeffizienten bezeichnet. Das Kanalaufmerksamkeitsmodul kann unter Verwendung von Gleichung wie folgt ausgedrückt werden:

$$CA(F) = \delta(MLP(AvgPool(F)) + MLP(MaxPool(F)))$$

wobei δ die sigmod-Aktivierungsfunktion bezeichnet.

Räumliches Aufmerksamkeitsmodul (SA): Räumliche Aufmerksamkeitskarten werden durch Extraktion räumlicher Zuordnungsbeziehungen erstellt. Es wurde festgestellt, dass eine einheitliche Verteilung der räumlichen Aufmerksamkeitsgewichte für jeden Kanal in der räumlichen Aufmerksamkeitskarte vermieden werden sollte, und die räumlichen Aufmerksamkeitsgewichte sollten dynamisch auf der Grundlage der spezifischen Merkmale jedes Kanals angepasst werden. Bild 3 veranschaulicht die Verwendung eines in der Tiefe trennbaren Faltungsmoduls zur Erfassung der räumlichen Beziehungen zwischen den Merkmalen, wodurch sichergestellt wird, dass die Beziehungen zwischen den Kanälen beibehalten werden und gleichzeitig die Rechenkomplexität verringert wird. Eine Multiskalenstruktur wird verwendet, um die Fähigkeit der Faltungsoperation zu verbessern, räumliche Beziehungen zu erfassen, räumliche Aufmerksamkeitsmerkmalskarten unter verschiedenen Kanälen zu erzeugen und schließlich die 1×1-Faltungsfusion der räumlichen Aufmerksamkeitsmodule verschiedener Kanäle in einem Hybridmodul zu verwenden, um optimierte Merkmalskarten zu erzeugen.

$$SA(F) = Conv_{1\times1}(\sum_{i=0}^{3} Branch_i(DwConv(F)))$$

wobei DwConv die tiefen-separierbare Faltungsoperation bezeichnet, $Branch_i$, $i \in \{0,1,2,3\}$ die Faltungsoperation des i-ten Zweigs, und zur Erfassung von Mehrskaleninformationen werden zwei kleine Kernel-Faltungen verwendet, um die Kanalfaltung des großen Kernels anzunähern, und es werden Faltungskerne unterschiedlicher Größe festgelegt.

Die Modelle der Kanalaufmerksamkeit und der räumlichen Aufmerksamkeit konzentrieren sich auf die semantischen bzw. räumlichen Informationen der Merkmale, und die experimentellen Ergebnisse zeigen, dass die sequentielle Anordnung der beiden Aufmerksamkeitsmodule effektiver ist als die parallele Anordnung.

Bei der Methode zur Verbesserung von Mischdaten werden zunächst zwei verschiedene Trainingsdaten (d_i, l_i) und (d_j, l_j) zufällig aus dem Trainingssatz ausgewählt, wobei d die Trainingsdaten und l die gelabelten Daten bezeichnet und $i \neq j$. Dann werden diese beiden Zufallsstichproben durch den folgenden Algorithmus berechnet, um neue Probendaten zu erzeugen:

$$\hat{d} = \lambda d_i + (1-\lambda) d_j \;;$$

$$\hat{l} = \lambda l_i + (1-\lambda) l_j \;;$$

Dabei ist $\lambda \in [0,1]$ ein Hyperparameter, der der $Beta \in (\alpha, \alpha)$ Verteilung gehorcht, d.h. der Intensitätskoeffizient, der die Interpolation steuert. Die Größe von α ist proportional zur Intensität der Interpolation, die die Mischtiefe der Bildmusterdaten beeinflusst, und wenn G=0 ist, werden die Daten nicht gemischt und Mixup funktioniert nicht.

Im Experiment werden zunächst N Stichproben aus dem Originaldatensatz ausgewählt, um

den Stichprobendatensatz zu bilden, und der N-Stichproben-Datensatz wird dem Zufallsumsturz unterzogen, ausgedrückt als Gleichung *.

$$S_{\mathrm{p}}=\{(d'_{P1},l'_{P1}),(d'_{P2},l'_{P2}),\ldots(d'_{PN},l'_{PN})\}$$

Die N Stichproben werden dann wie in Gleichung * ausgedrückt erneut randomisiert, was zu zwei zufällig angeordneten Stichproben-Datensätzen führt.

$$S_{\mathrm{q}}=\{(d'_{q1},l'_{q1}),(d'_{q2},l'_{q2}),\ldots(d'_{qN},l'_{qN})\}$$

Zu diesem Zeitpunkt sind die Stichproben, die den beiden Stichprobensätzen S_{p} und S_{q} entsprechen, gemäß der folgenden Gleichung unterschiedlich:

$$(\mathrm{x}'_{pi},\mathrm{y}'_{pi})\neq(\mathrm{x}'_{qi},\mathrm{y}'_{qi})$$

Unter Verwendung von Gleichung (1) und Gleichung (2) für die S_{p}- und S_{q}-Stichproben ergibt sich ein neuer Stichproben-Datensatz, S_{mixup}, der wie folgt ausgedrückt werden kann:

$$S_{\mathrm{mix}up}=\{(\tilde{d}_1,\tilde{l}_1),(\tilde{d}_2,\tilde{l}_2),\ldots(\tilde{d}_N,\tilde{l}_N)\}$$

Der neue Datensatz nach dem Mischen durch Mixup ersetzt den Originaldatensatz für das Training.

CIoU wurde als Regressionsverlustfunktion verwendet, wie in der folgenden Gleichung ausgedrückt:

$$L_{loc}=1-IoUI\ (B,\ B_{\mathrm{gt}})+\frac{d^2}{c^2}+\alpha v\ ;$$

$$\alpha=\frac{v}{1-IoU+v}\ ;$$

$$v=\frac{4}{\pi^2}(\arctan,\frac{w^{gt}}{h^{gt}}-\arctan\frac{w}{h})^2\ ;$$

Dabei steht w für die Breite des Vorhersagekastens, h für die Höhe des Vorhersagekastens, w^{gt} für die Breite des Zielkastens und h^{gt} für die Höhe der Zielbibliothek.

Ausführungsform 2

Auf der Grundlage von Ausführungsform 1 wird Folgendes vorgeschlagen: Das Verfahren zur Erstellung des in den Experimenten dieser Studie verwendeten Datensatzes der Zellproben aus dem Gebärmutterhals wurde wie folgt durchgeführt: Das Präparat der Zellen aus dem Gebärmutterhals wurde mit einem vollautomatischen Scanner für Zellen aus dem Gebärmutterhals gescannt, der von JingCast Biologicals selbst entwickelt wurde, und ein Panorama der Zellproben aus dem Gebärmutterhals wurde erstellt. Dann wurde das Panorama der Proben gescreent, und es wurden 150 positive und 150 negative Proben ausgewählt. Von den positiven Proben enthielten 50 abnorme Zellhaufen; die negativen Proben enthielten jeweils 15 Mykobakterien, Trichomonaden, Aktinomyzeten, Clue-Zellen, Herpes und Aktinomyzeten; die restlichen 60 negativen Proben

waren frei von mikrobiellen Infektionen und enthielten 5 Proben mit atrophischen Zellen und 5 Proben mit Chemotaxis-Zellen. Ein Pathologe wurde dann gebeten, die Panoramaproben zu beschriften und die beschrifteten Diagramme zu schneiden, um den Datensatz der auf Gebärmutterhalsflüssigkeit basierenden Zellproben zu erhalten.

Die Beispieldaten werden durch Mosaikdatenverbesserung, affine Transformation und Perspektive-Transformation verbessert, auf deren Grundlage das Mixup-Modul zur Datenverbesserung hinzugefügt wird. Die Kernidee besteht darin, zwei Bilder aus jedem Stapel nach dem Zufallsprinzip auszuwählen und sie in einem bestimmten Verhältnis zu mischen, um ein neues Bild zu erzeugen, und der gesamte Trainingsprozess verwendet das gemischte neue Bild zum Training, wobei das Originalbild nicht mehr in das Training einbezogen wird.

In diesem Experiment wurde der Zelldatensatz aus der Gebärmutterhalsflüssigkeit bei der Durchführung des Tests in insgesamt 10 Kategorien eingeteilt. Die einzelnen Kategorien sind in Tabelle 1 aufgeführt.

Tabelle 1 Datensatz von Zellen auf der Basis von Gebärmutterhalsflüssigkeit

Zelltyp	Anzahl	Gattung	Zelltyp	Anzahl	Gattung
LSIL	2657	Positiv	PARASITE	2305	Mikroorganismen
HSIL	2019	Positiv	INDICATORS	2085	Mikroorganismen
AGC	1003	Positiv	LEPTOTRICHIA	238	Mikroorganismen
ENDOMETRIAL	991	Positiv	ACTINOMYCES	775	Mikroorganismen
FUNGAL	2005	Mikroorganismen	HERPES	533	Mikroorganismen

Dieses experimentelle Training wurde mit dem Betriebssystem Ubuntu18.04 unter Verwendung von Pytorch1.8.1, dem CUDA11.4 Deep-Learning-Trainingsframework und der NIVDIA RTX3090TiGPU durchgeführt, um das Modelltraining zu beschleunigen, und die Trainingsparameter wurden wie folgt festgelegt: Epoch ist auf 100 mal eingestellt, batch_size ist auf 32 eingestellt, die Eingabebildgröße ist 1024x1024, und die Kombination aus regulärem Training und Freeze-Training wird angenommen, um während des Trainingsprozesses zu iterieren, um das Overfitting-Phänomen, das während des Trainingsprozesses auftreten kann, weiter zu hemmen, und das spezifische Trainingsschema ist wie folgt:

Schritt 1: Konventionelle Trainingsmethode, zufällige Auswahl der Hälfte der Daten im Datensatz, Durchführung eines konventionellen Trainings für das ausgewählte Netzwerk, alle Parameter werden in diesem Prozess gleichzeitig aktualisiert, und man erhält Modell1;

Schritt 2: Auswahl des besten Modells im ersten Schritt des konventionellen Trainings, zufällige Auswahl der Hälfte der Trainingsdaten, mit 100 Epochen als Stufe, die ersten 40 Epochen werden die Kopfschicht eingefroren, und nur die Backbone-Schicht wird trainiert; In der 40. bis 100. Epoche wird das Backbone eingefroren und nur die HEAD-Schicht trainiert, und die Parameter der Backbone- und HEAD-Schicht werden nacheinander aktualisiert, um Modell2 zu erhalten;

Schritt 3: Wählen Sie auf der Grundlage von Modell2 alle Trainingsdaten aus, passen Sie die Lernrate um eine Größenordnung nach unten an und trainieren Sie in eingefrorener Weise gemäß

Schritt 2, um das endgültige Modell zu erhalten.

Die gängigen Bewertungsmaßstäbe, die derzeit für Zielerkennungsmodelle verwendet werden, sind Genauigkeit, Präzision, Wiedererkennungswert und mAP, wobei der Wiedererkennungswert verwendet wird, um zu bewerten, ob die Vorhersage des Netzwerkmodells umfassend ist oder nicht, und die Präzision, um zu beurteilen, ob die Vorhersageergebnisse genau sind oder nicht. AP misst, wie gut das trainierte Modell in jeder Kategorie ist, und mAP misst, wie gut das trainierte Modell in allen Kategorien ist; die oben genannten Metriken werden mit der folgenden Formel ausgedrückt

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

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

$$\mathrm{m}AP = \frac{\sum_{n=1}^{N} AP(n)}{N}$$

wobei $p(r)$ die Precision-Recall-Kurve bezeichnet, N die Gesamtzahl der Kategorien der kategorisierten Daten ist und n die Kategorie der aktuellen Daten ist.

Obwohl Ausführungsformen der vorliegenden Erfindung gezeigt und beschrieben worden sind, wird der Fachmann erkennen, dass eine Vielzahl von Änderungen, Modifikationen, Substitutionen und Variationen an diesen Ausführungsformen vorgenommen werden können, ohne von dem Prinzip und dem Geist der vorliegenden Erfindung abzuweichen, deren Umfang durch die beigefügten Ansprüche und deren Äquivalente begrenzt ist.

Ansprüche

1. Ein Verfahren zur automatischen Klassifizierung und Erkennung von Zellen auf der Basis von Gebärmutterhalsflüssigkeit, dadurch gekennzeichnet, dass das Verfahren die folgenden Schritte umfasst:

Konstruktion eines CPCA-Mechanismus für räumliche Aufmerksamkeit unter Verwendung eines tiefen Faltungsmoduls zur Erzeugung einer dynamisch verteilten räumlichen Aufmerksamkeitskarte auf jedem Kanal;

Um die Zielvielfalt zu erhöhen und das Problem unzureichender Datensätze und unausgewogener Stichproben für die Erkennung kleiner Ziele anzugehen, zusätzlich zur Mosaik-Datenverbesserung, affinen Transformation und perspektivischen Transformation, Hinzufügen des Mixup-Moduls zur Datenverbesserung, d.h. zufällige Auswahl von zwei Bildern aus jedem Batch und Mischen dieser Bilder in einem bestimmten Verhältnis, um ein neues Bild zu erzeugen;

Verbesserung der Verlustfunktion durch Ersetzen der Verlustfunktion „Generalised Intersection and Combination Ratio“ GIoU durch die Verlustfunktion „Complete Intersection and Combination Ratio“ CIoU, um die Leistung der Bounding-Box-Regression und der Lokalisierung zur Identifizierung von zervikalen pathologischen Zellfraktionen mit komplexer Umgebung zu verbessern.

2. Ein Verfahren zur automatischen Klassifizierung und Erkennung von Zellen auf der Basis von Gebärmutterhalsflüssigkeit nach Anspruch 1, dadurch gekennzeichnet, dass das CPCA-Modul dem Backbone-Netzwerk und dem Halsnetzwerk hinzugefügt und nach dem C3-Modul platziert wird, und dass das CPCA-Modul im Prozess des Netzwerktrainings das Lernen der Merkmalsinformationen des kleinen Ziels durch dynamisches Einstellen der Kanal- und Raumaufmerksamkeitsmodul-Gewichte verbessert.

3. Ein Verfahren zur automatischen Klassifizierung und Erkennung von Zellen auf der Basis von Gebärmutterhalsflüssigkeit nach Anspruch 1, dadurch gekennzeichnet, dass: der Mechanismus der Kanal-A-priori-Faltungsaufmerksamkeit CPCA nacheinander zwei Modulaufgaben der Kanalaufmerksamkeit CA und der räumlichen Aufmerksamkeit SA durchführt, wobei angenommen wird, dass das Eingangsbildmerkmal $F \in R^{C\times H\times W}$ ist. Eine eindimensionale Kanalaufmerksamkeitsmerkmalskarte $M_C \in R^{C\times 1\times 1}$ wird zunächst durch das Kanalaufmerksamkeitsmodul CA erhalten, und dann wird M_C Element für Element mit dem Eingangsmerkmal F multipliziert, um den Kanalaufmerksamkeitswert aus der räumlichen Dimension zu übertragen, um so ein besseres Bildmerkmal $F_C \in R^{C\times H\times W}$ mit Kanalaufmerksamkeit zu erhalten; Das räumliche Aufmerksamkeitsmodul SA erzeugt eine räumliche 3D-Aufmerksamkeitskarte $M_s \in R^{C\times H\times W}$ durch Datenverarbeitung der Bildmerkmale F_C, und gibt dann schließlich die verarbeiteten Bildmerkmale $\hat{F} \in R^{C\times H\times W}$ durch elementweise Multiplikation von M_s und F_C aus, und der gesamte Prozess wird durch die folgende Gleichung ausgedrückt:

$$F_C = CA(F) \otimes F$$

$$\hat{F} = SA(F_C) \otimes F_C$$

4. Ein Verfahren zur automatischen Klassifizierung und Erkennung von Zellen auf der Basis von Gebärmutterhalsflüssigkeit nach Anspruch 3, dadurch gekennzeichnet, dass das Kanalaufmerksamkeitsmodul eine Kanalaufmerksamkeitsmerkmalskarte erzeugt, indem es die in den Bildmerkmalen vorhandenen Beziehungen zwischen den Kanälen untersucht, und dass die räumliche Aufmerksamkeitskarte durch Extraktion räumlicher Abbildungsbeziehungen erzeugt wird.

5. Ein Verfahren zur automatischen Klassifizierung und Erkennung von Zellen auf der Basis von Gebärmutterhalsflüssigkeit nach Anspruch 1, dadurch gekennzeichnet, dass das Mischmodul ein lineares Differenzverfahren anwendet, um zwei unterschiedliche Trainingsproben und Etikettenproben aus dem ursprünglichen Datensatz zum Mischen und Erzeugen einer neuen Probe auszuwählen, und dass zwei unterschiedliche Trainingsprobendaten (d_i, l_i) und (d_j, l_j) zufällig aus dem Trainingssatz ausgewählt werden. Wobei d die Trainingsdaten, l die etikettierten Daten und $\mathrm{i} \neq j$ bezeichnet, dann werden die beiden Zufallsstichproben durch den folgenden Algorithmus berechnet, um eine neue Stichprobe zu erzeugen, der Algorithmus ist in der folgenden Gleichung angegeben:

$$\hat{\mathrm{d}} = \lambda \mathrm{d}_{\mathrm{i}} + (1-\lambda) d_j \ ;$$

$$\hat{l} = \lambda l_{\mathrm{i}} + (1-\lambda) l_j \ ;$$

Dabei ist $\lambda \in [0,1]$ ein Hyperparameter, der der $Beta \in (\alpha, \alpha)$ Verteilung gehorcht, d.h. der Intensitätskoeffizient, der die Interpolation steuert. Die Größe von α ist proportional zur Intensität der Interpolation, die die Mischtiefe der Bildmusterdaten beeinflusst, und wenn G=0 ist, werden die Daten nicht gemischt und Mixup funktioniert nicht.

6. automatisches Zellsortier- und -nachweisverfahren auf der Basis von Zervixflüssigkeit nach Anspruch 1, dadurch gekennzeichnet, dass CIoU als Regressionsverlustfunktion verwendet wird, die in der folgenden Formel ausgedrückt wird:

$$L_{loc} = 1 - IoUI\ (B,\ B_{\mathrm{gt}}) + \frac{d^2}{c^2} + \alpha v$$

$$\alpha = \frac{v}{1 - IoU + v}$$

$$v = \frac{4}{\pi^2}(\arctan, \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h})^2$$

Dabei steht w für die Breite des Vorhersagekastens, h für die Höhe des Vorhersagekastens, w^{gt} für die Breite des Zielkastens und h^{gt} für die Höhe der Zielbibliothek.

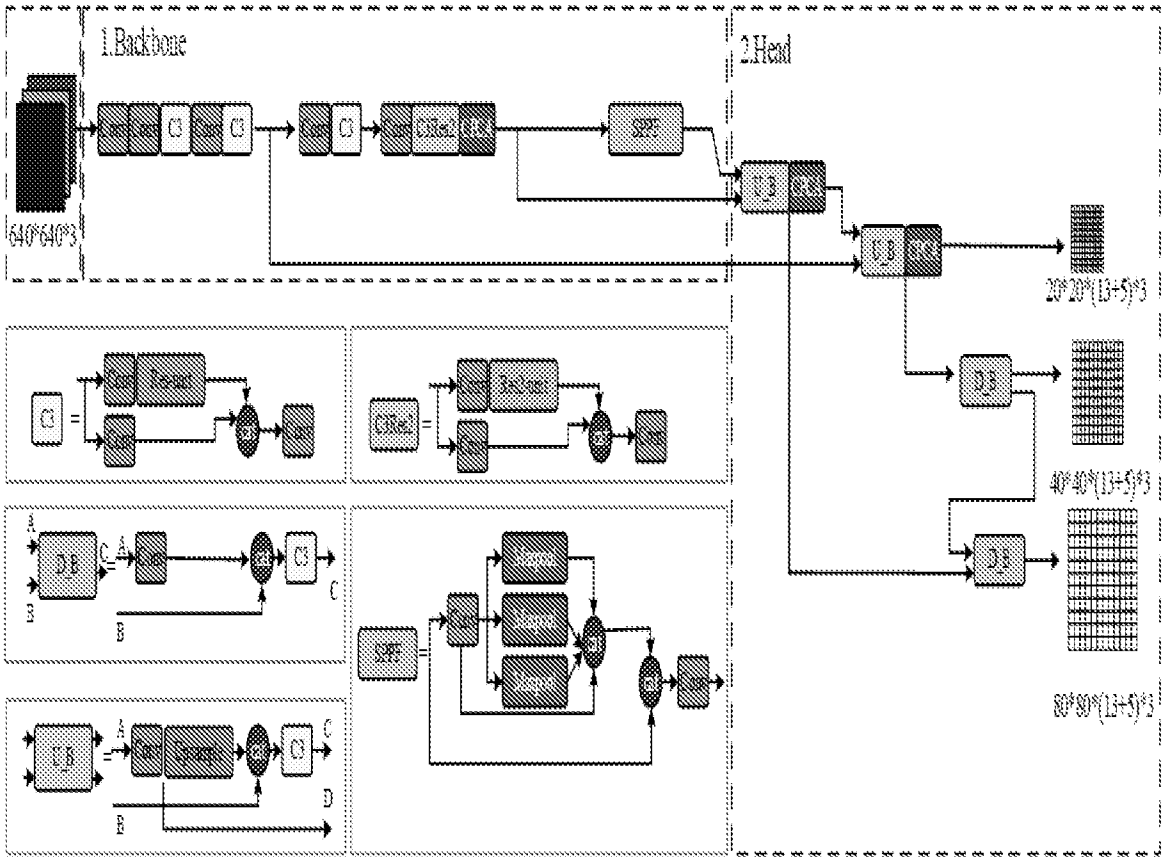

Bild 1

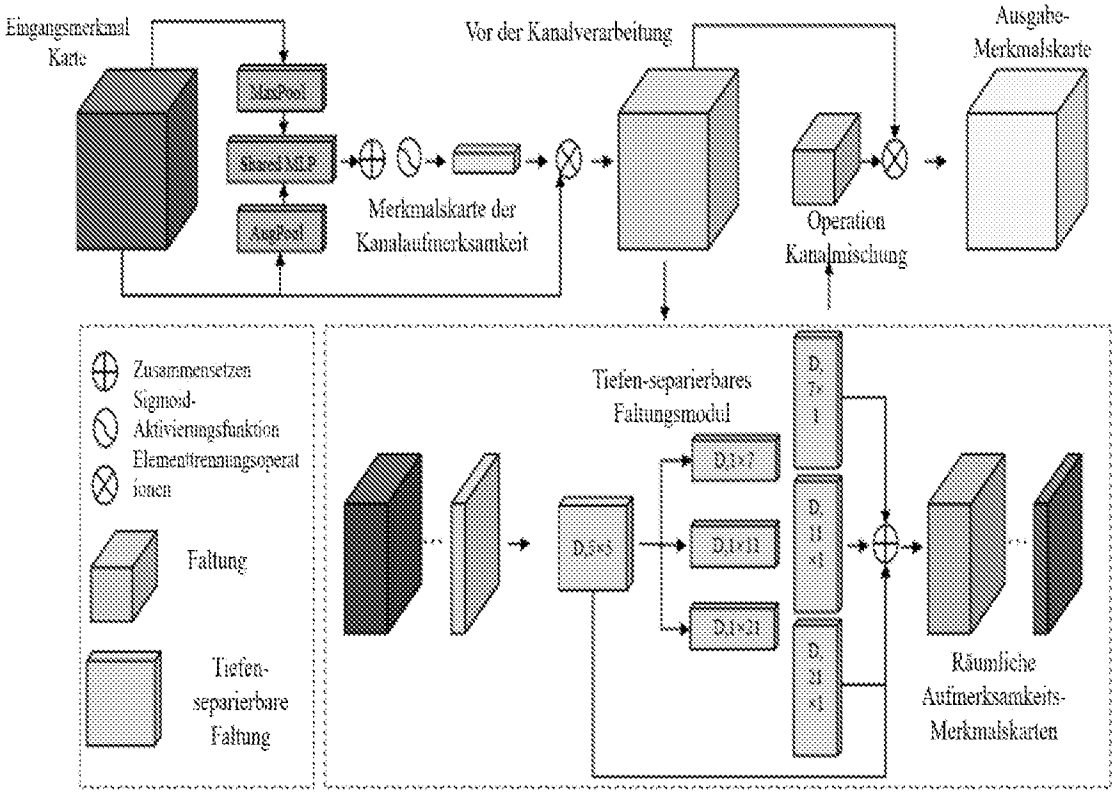
Eingangsmerkmal Karte
Vor der Kanalverarbeitung
Ausgabe-Merkmalskarte
Merkmalskarte der Kanalaufmerksamkeit
Operation Kanalmischung
Zusammensetzen
Sigmoid-Aktivierungsfunktion
Elementtrennungsoperationen
Faltung
Tiefen-separierbare Faltung
Tiefen-separierbares Faltungsmodul
Räumliche Aufmerksamkeits-Merkmalskarten

Bild 2