استعمال مستويات التصنيف الثلاثية لتعزيز تأثير الشبكات العصبية المعيارية

أ .م. د مهدي وهاب نصرالله[١] احمد تركي عبد علي[٢]

جامعة كربلاء المقدسة- كلية الادارة والاقتصاد-قسم الاحصاء

ahmed.turki@uokerbala.edu.iq[2]

mehdi.wahab@uokerbala.edu.iq[1]

المستخلص: يحتوي النظام البصري البشري على تسلسل هرمي للوحدات النمطية التي تشارك في الإدراك البصري عند مستويات التصنيف الفائقة والأساسية والفرعية خلال العقود السابقة ، تم اقتراح انموذج حسابي مختلف الى المعالجة الهرمية للتغذية الأمامية للقشرة البصرية ، ولكن تم تجاهل العديد من الخصائص الهامة للنظام البصري ، مثل آليات التعلم والمعالجة العصبية الفعلية. نقترح نموذجًا حسابيًا للتعرف على الكائنات في مستويات تصنيف مختلفة ، حيث يتم استعمال شبكة عصبية متصاعدة مجهزة بقاعدة التعلم المعزز كوحدة نمطية في كل مستوى تصنيف. تعمل كل وحدة على حل مشكلة التعرف على الكائن في كل مستوى تصنيف ، بناءً على أول ارتفاع في الخلايا العصبية الخاصة بالفئة في الطبقة الأخيرة ، دون استعمال أي مصنف خارجي. وفقًا للمعلومات المطلوبة في كل مستوى تصنيف ، يتم استعمال الصور ذات تمرير النطاق التصفية. يتم تقييم أداء نموذجنا المقترح من خلال معايير تقييم مختلفة مع ثلاث مجموعات بيانات معيارية وتحقق تحسن كبير في دقة التعرف على نموذجنا المقترح في جميع التجارب

الكلمات المفتاحية: مستويات التصنيف ، الانموذج الحسابي، المعالجة من أسفل إلى أعلى ، الشبكات العصبية .

Abstract: The human visual system contains a hierarchy of modules that participate in visual perception at hyperboloid, primary and sub-classification levels. During previous decades, a different computational model has been proposed to the hierarchical treatment of the anterior cortical feeding of the visual cortex, but many important characteristics of the visual system have been ignored, such as Actual Neuronal Learning and Processing Mechanisms. We propose a mathematical model to recognize objects at different classification levels, Where an escalating neural network equipped with the reinforcement learning base is used as a modular unit at each classification level. Each unit solves the object recognition problem at each classification level, based on the first class-specific neuron spike in the last layer, without using any external classifier. According to the information required at each classification level, the filter band-pass images are used. The performance of our proposed model is evaluated through various evaluation criteria with three standard data sets and a significant improvement in the accuracy of our proposed model recognition was achieved across all trials.

١. المقدمة:

يمكن التعرف على مستويات مختلفة من التحليل الهرمي. قام علماء عمومًا بتقسيم هذا التجربة إلى ثلاثة مستويات:

المستوى الأعلى :على سبيل المثال (الحيوان مقابل غير الحيوان)

والمستوى الأساسي: على سبيل المثال (الطيور مقابل الثدييات)

والثالث المرؤوس: على سبيل المثال (الحمام مقابل البط) (L. Kauffmann :2015) .

وجدت بعض الدراسات دليلاً على أن مصطلحات التعريف هذه تقع ضمن تصنيف المستويات[L. Kauffmann (2014) ، K. Fukushima (2003)].

يكشف الدليل البيولوجي أن المدخلات المرئية تتم معالجتها من(الترددات المكانية المنخفضة) إلى السمات الدقيقة (الترددات المكانية العالية) [S. Grossberg :2007 ، T. Serre :2007]. بشكل عام تتم معالجة عناصر الترددات الأكبر والعالية والمنخفضة (LSF) للمدخلات بأولوية أعلى من عناصر التردد المكاني الصغير والمحلي والعالي (HSF) [K. Fukushima :2003 , A.Yuille :2016] خاصةً عندما تكون مستويات التصنيف مختلفة. أظهرت نتائج بعض تجارب علم النفس الفيزيائي والانموذج الحسابي أن تصنيف المستوى الأعلى يعتمد بشكل أساسي على LSFs وأن دقة المستويات الأساسية والثانوية تتعزز بزيادة التردد المكاني .تم فحص الانموذج حسابية المختلفة لكل مستوى تصنيف على حدة. تم اقتراح هذه النماذج للتحليل الهرمية للتغذية الأمامية لمهمة التعرف على الكائن في سلسلة من الأسفل إلى الأعلى من المناطق القشرية البشرية [T. Serre :2007]. على الرغم من أن هياكل هذه النماذج مستوحاة من القشرة البصرية للإنسان (التسلسل الهرمي للطبقات ذات الحقول الاستباقية المتزايدة تدريجيًا) ، يستمر البحث عن النماذج الحسابية المستوحاة من الدماغ ويجذب المزيد والمزيد من الباحثين من جميع أنحاء العالم. بشكل عام ترتبط التحديات البحثية لمهمة التعرف على الأشياء في مستويات التصنيف المختلفة بشكل أساسي بالمسائل الخمس التالية.

ISSN:2618-0278 Vol. 3 No. 6 June 2021

أولاً: حتى الآن لم يتم اقتراح أي انموذج حسابي لأداء مهمة التعرف على مستويات التصنيف الثلاثة ، أي أن النماذج الحسابية الحالية تتعامل مع كل مستوى تصنيف على حدة.

ثانيًا: يزداد تعقيد التعرف على الأشياء بالانتقال من المستوى الأعلى إلى المستوى الثانوي.

ثالثًا: معظم النماذج الحسابية الحالية ليست معقولة حقًا من الناحية البيولوجية لأن المعالجة العصبية الفعلية وآليات التعلم للقشرة البصرية مهملة في هذه النماذج.

رابعًا: أثناء العديد من مهام التعرف على الكائنات قد يكون الإدخال غامضًا ، أي يظهر مع اختلافات مختلفة أو ضوضاء أو انسداد ، وبالتالي ، يتم الوصول إلى معلومات جزئية ومحدودة فقط حول الكائنات مما يعيق التعيين الواضح من الإدخال إلى فئة الكائن.

ان التقدم في مستويات التصنيف ، يزداد تعقيد تدريجيًا ويصبح التمييز الدقيق بين الأشياء المتشابهة أكثر صعوبة [S. Grossberg :2007C. A. Collin :2005،] ، خاصة على المستوى الثانوي. يعد التعرف على الفئات الدقيقة على المستوى الثانوي أمرًا صعبًا للغاية لأن الاختلافات المرئية بين الفئات الدقيقة صغيرة ويمكن أن يطغى عليها تنوع الكائنات بسهولة. ترجع الصعوبات في هذه المهمة إلى توطين المنطقة التمييزية والتعلم الدقيق للميزات [T. Serre :2007]. وبالتالي ، قد لا تكون المناطق التي يحددها الإنسان أو المناطق التي تم تعلمها من خلال الأساليب الحالية هي الأمثل لتصنيف الآلة وأيضًا ، من الصعب تعلم الاختلافات المرئية الدقيقة الموجودة في المناطق المحلية من الفئات الدقيقة المتشابهة.

ترسل وحدات الحوسبة في معظم النماذج الحسابية ، مثل الشبكة العصبية لالتواء غامض (DCNN) ، والتي تتوافق مع مستوى التنشيط ، بينما تتواصل الخلايا العصبية البيولوجية مع بعضها البعض عن طريق إرسال النبضات. تطلق الخلايا العصبية ارتفاعًا فقط عندما يتعين عليها نقل رسالة مهمة ، ويمكن تشفير بعض المعلومات في أوقات ارتفاعها. مستوحاة من النماذج الواقعية بيولوجيًا للخلايا العصبية ، تم تطوير الشبكات العصبية المتصاعدة (SNN) لإجراء العمليات الحسابية. درس العديد من الباحثين معلمات تكوين SNN ، مثل عدد الطبقات (الضحلة أو العميقة) [L. Kauffmann :2014، A.Yuille :2016، K. Fukushima :2003] ، وأنواع مخططات الاتصال (المتكررة والمتصلة بالكامل) ، يعمل STDP جيدًا في العثور على ميزات متكررة إحصائيًا مثل أي قاعدة تعلم غير خاضعة للإشراف ، فإنه يواجه صعوبات في اتخاذ القرار ، وبالتالي ، عادةً ما تكون المصنفات الخارجية مطلوبة. لذلك ، تقترح العديد من الدراسات استعمال التعلم المعزز (RL) الذي له تأثير كبير على صنع القرار وتكوين السلوك. في RL ، يتم تشجيع المتعلم على تكرار السلوكيات المجزية وتجنب تلك التي تؤدي إلى العقوبات [C. A. Collin :2005، K. Fukushima :2003]. تم استلهام RL من نظام المكافأة في الدماغ في حين أن التعلم غير الخاضع للإشراف هو أحد القوى المحركة للمرونة ، فإن STDP يهمل المعلومات المتعلقة بالثواب والعقاب قامت العديد من الدراسات بنمذجة دور نظام المكافأة عن طريق تعديل تغيير الوزن المحدد بواسطة STDP ، والذي يسمى STDP المعدلة في وقت تلقي إشارة المكافأة أو العقوبة المعدلة.

2.1 المواد والطرق الاحصائية:

في هذا البحث نقترح نموذجًا حسابيًا جديدًا للتعرف على الأشياء في مستويات تصنيف مختلفة ، وهي فوقية ، وأساسية ، ومرؤوسة ، تحاكي المعالجة الهرمية في القشرة البصرية. لكل مستوى تصنيف ، يتم استعمال وحدة مقابلة في النموذج. تشتمل كل وحدة على شبكة عصبية هرمية متدرجة مزودة بقاعدة تعلم R-STDP ، بعنوان R-SNN ، لحل مهمة التعرف على الكائن في مستوى التصنيف هذا. يتم التعرف في كل وحدة على أساس النشاط العصبي لطبقتها الأخيرة ، دون استعمال أي مصنف خارجي.

باستعمال جميع نطاقات التردد لصورة الإدخال في كل مستوى تصنيف ، يتم إدخال معلومات إضافية وفائضة في الوحدات النمطية. فإن دور تطبيق المعلومات ذات الصلة لكل مستوى يتلاشى وفقًا للمعلومات المطلوبة في كل مستوى تصنيف ، يتم استعمال الصور ذات الصلة بمرشح تمرير النطاق. وبشكل أكثر تحديدًا ، يتم استعمال معلومات الترددات المكانية المنخفضة والمتوسطة والعالية (LSF و ISF و HSF) للصور المدخلة في الوحدة الأولى والثانية والثالثة من النموذج المقترح ، على التوالي. [L. Kauffmann :2014، C. A. Collin :2005، K. Fukushima :2003].

٢,٢ بناء الهيكل العام للوحدات:

التخطيطي الهيكل كل وحدة مع طبقتين بسيطتين وطبقتين معقدتين يتم ترتيبهما بالتناوب ، ويجب تكييف معلمات كل وحدة لمهمة التعرف المطلوبة.

الطبقة الأولى البسيطة (c_1) ، تستخرج الحواف الموجهة من صورة الإدخال التي تمت تصفيتها بالتردد وتحولها إلى زمن انتقال سريع باستعمال تشفير الكثافة إلى زمن الوصول. تحقيقا لهذه الغاية ، يتم تحويل المدخلات مع مرشحات Gabor من أربعة اتجاهات مختلفة يتم تحقيق أربع خرائط معالم في هذه الطبقة ، يمثل كل منها الحواف في اتجاه معين. لكل خريطة معالم ، يتم وضع شبكة ثنائية الأبعاد بنفس الحجم تحتوي على خلايا عصبية وهمية لانتشار التموج. يتم تحويل خرائط الميزات التي تم الحصول عليها إلى زمن انتقال الارتفاع ، والذي يتناسب عكسياً مع بروز الحواف ، أي كلما كانت الحافة أكثر بروزًا ، كلما تم نشر الارتفاع المقابل في وقت مبكر. يتم فرز بترتيب تصاعدي حسب أزمنة انتقالها ويتم نشرها بالتتابع (أي يتم نشر الارتفاع الأول في الخطوة الزمنية t = 1 ، والثاني في t = 2 ، إلخ). [L. Kauffmann :2014، A.Yuille :2016، T. Serre :2007]

ان هيكل وحدة R-SNN بأربع طبقات يتم استخراج حد الصورة المدخلة في طبقة S_1. يتم تنفيذ عملية التجميع المحلي الأقصى في طبقة C_1 ويتم نشر النتوءات في الطبقة S_2 ، توجد شبكات ثنائية الأبعاد من الخلايا العصبية IF ، والتي تتعلم استخراج الميزات المعقدة باستعمال قاعدة التعلم R-STDP. يتم تحديد القرار النهائي لوحدة R-SNN بواسطة الخلايا العصبية في الطبقة C_2 المرتبطة بتسميات الفئات.

ISSN:2618-0278 Vol. 3 No. 6 June 2021

تقوم الطبقة الثانية (C_1) بتجميع محلي على التموجات القادمة من S_1. في طبقة C_1 ، يتم وضع أربعة مشابك عصبية ثنائية الأبعاد تتوافق مع كل اتجاه. ينفذ كل خلية عصبية في C_1 عملية تجميع محلية بحجم $w_{c1} \times w_{c1}$ و ($s_{c1} = w_{c1}$-1) s_{c1} على الخلايا العصبية S_1 في الشبكة المقابلة . تجمع الطبقة الثالثة (S_2) المعلومات الواردة من C_1 حول الحواف الموجهة وتحولها إلى ميزات أكثر تعقيدًا. تتكون هذه الطبقة من شبكات n 2-D من الخلايا العصبية المدمجة وإطلاق التاثير (IF) مع حد θ. لا يوجد تسرب لهذه الخلايا العصبية ويُسمح لها بالتصوير مرة واحدة على الأكثر أثناء عرض الصورة[K. Fukushima :2003 ' L. Kauffmann :2015].

يتم استعمال آلية تقاسم الوزن للخلايا العصبية في نفس الشبكة. يتم إنشاء الأوزان الأولية بشكل عشوائي. يتم تحديث إمكانات غشاء الخلايا العصبية i^{th} في S2 في الخطوة الزمنية t بالمعادلة التالية:

$$V_i(t) = V_i(t-1) + \Sigma_{j \in P(i)} W_{i,j} \times \delta(t - t_j) \qquad \ldots(1)$$

حيث **(Pi)** هي الخلايا العصبية ، **tj** هي وقت إطلاق الخلية العصبية j^{th} **في** C_1 ، **δ** هي وظيفة دلتا Kronecker ، و Wi تحدد الوزن المشبكي بين الخلايا العصبية i و j. عندما تطلق الخلايا العصبية S_2 ، يتم تحديد أوزانها المشبكية وفقًا لترتيب الارتفاعات قبل وبعد التشابك العصبي ، بالإضافة إلى إشارة المكافأة / العقوبة. هذه الإشارة مشتقة من نشاط الطبقة التالية ، والتي تشير أيضًا إلى قرار الوحدة. [A.Yuille :2016 ' C. A. Collin :2005]

تحدد الطبقة الرابعة (C_2) القرار النهائي للوحدة. يستعمل هذا القرار لتوليد إشارة مكافأة / عقاب ، والتي تعدل على التشابكية للخلايا العصبية S_2. تحتوي هذه الطبقة على n خلايا عصبية ، خلية عصبية واحدة لكل شبكة في طبقة S_2 ، وتقوم كل خلية عصبية بعملية تجميع شاملة. يتم حساب وقت إطلاق الخلايا العصبية i^{th} (i=1,...,n) في طبقة C_2 على النحو التالي[L. Kauffmann :2014]:

$$ti = \min_{j \in P(i)} \{tj\} \qquad \ldots(2)$$

حيث يشير (Pi) إلى الخلايا العصبية S_2 في الشبكة العصبية i^{th} (الخلايا العصبية قبل التشابك) ، و هو وقت إطلاق الخلايا العصبية j^{th} في طبقة S_2. يشير نشاط الخلايا العصبية C_2 إلى قرار الوحدة. لذلك ، يتم تقسيم الخلايا العصبية C_2 إلى عدة مجموعات ويتم تعيين كل مجموعة إلى فئة معينة من صورة الإدخال في تلك الوحدة. بهذه الطريقة ، تحدد المجموعة التي تنشر أول ارتفاع بين مجموعات C_2 قرار الوحدة. لنفترض أن n تشير إلى عدد شبكات S_2 بينما تشير m إلى عدد فئات الإدخال. ثم يتم حساب قرار الوحدة (D) لكل صورة إدخال من خلال [L. Kauffmann :2015]

$$D=g(F), \qquad F=\min_{1 \le i \le n} \{ti\} \qquad \ldots(3)$$

حيث يشير F إلى الخلية العصبية C_2 التي تنشط أولاً ، هي وقت إطلاق الخلايا العصبية i^{th} في طبقة C_2 ، و $g: \{1,\ldots,n\} \rightarrow \{1,\ldots,m\}$ هي وظيفتها جدولة مجموعة من الخلايا العصبية C_2. إذا تطابق D (لا يتطابق) مع الفئة الصحيحة لصورة الإدخال ، تتلقى الوحدة إشارة مكافأة (عقاب). عندما ينطلق أكثر من خلية عصبية واحدة ، يتم تحديد الخلية ذات الحد الأدنى من وقت الارتفاع. علاوة على ذلك ، عندما تكون جميع الخلايا العصبية C2 صامتة ، لا يتم إنشاء إشارة المكافأة / العقوبة وبالتالي ، لا يتم تغيير الأوزان. [A.Yuille :2016]

٣,٢ دراسة وتصنيف للأوزان:

يتم تحديث أوزان الخلايا العصبية S_2 باستعمال قاعدة التعلم R-STDP. وفقًا لصحة / خطأ قرار الوحدة ، يتم ضبط حجم / أو تغيير الوزن بواسطة إشارة المكافأة / العقوبة ، حيث تكون الخلية العصبية ذات الارتفاع المبكر هي الفائزة وتكون مؤهلة لتحديث الأوزان المتشابكة ، والتي يتم مشاركتها بين الخلايا العصبية الأخرى في نفس الشبكة. في الواقع ، تحدد الخلايا العصبية الفائزة قرار الوحدة. وفقًا لـ R-STDP ، يتم حساب مقدار التغير في الوزن ($\Delta W_{i,j}$) للوصلة المشبكية بين عصبتين عصبيتين i، j من خلال:

$$\Delta W_{i,j} = \Gamma_{i,j} . W_{i,j} . (1 - W_{i,j}); \quad \Gamma i,j = \begin{cases} R.A_r.M_r^+ + P.A_p.M_p^- & : if\ t_j - t_i \le 0 \\ R.A_r.M_r^- + P.A_p.M_p^+ & : if\ t_j - t_i > 0 \end{cases} \qquad \ldots.(4)$$

حيث قيم M_r^+، M_r^- ، M_p^+ ، M_p^- تقيس حجم تغير الوزن ($M_r^+, M_p^+ > 0$ and $M_r^-, M_p^- < 0$). تعتمد القيمتان R و P على إشارة المكافأة / العقوبة المتولدة: إذا تم إنشاء إشارة المكافأة ، فإن R = 1 و P = 0 ؛ بينما إذا تم إنشاء إشارة العقوبة ، فإن R = 0 و P = 1. تحدث مشكلة فرط التخصيص عادةً بسبب التأثير غير المتوازن للمكافأة والعقاب في معدل مرتفع من العينات المصنفة بشكل صحيح وسوء التصنيف. في المعدل المرتفع لحالات التصنيف الصحيح ، يزداد معدل الحصول على المكافأة وتفضل الوحدة استبعاد العينات المصنفة بشكل خاطئ عن طريق الحصول على المزيد والمزيد من الانتقائية لتصحيح الحالات والبقاء صامتة بالنسبة للآخرين. وبالمثل ، في المعدل المرتفع لحالات التصنيف الخاطئ ، تتلقى الوحدة المزيد من إشارات العقاب ، مما يضعف الأوزان بسرعة ويولد خلايا عصبية ميتة أو انتقائية للغاية تغطي عددًا صغيرًا من المدخلات. للتعامل مع هذه الصعوبة ، يتم تطبيق عوامل الضبط ($A_r = Nincorrect / N$ ،$A_p = Ncorrect / N$) في تعديل الوزن ، حيث تشير Ncorrect و Nincorrect على التوالي إلى عدد العينات التي تم تصنيفها بشكل صحيح وغير صحيح خلال الدفعة الأخيرة من عينات الإدخال N. لذلك ، يتم موازنة تأثير عينات التدريب الصحيحة وغير الصحيحة على التجارب. وكذلك الضرب بين المصطلحات Γi و $W_{i,j} . (1 - W_{i,j})$ تؤدي إلى الحفاظ على الأوزان بين النطاق [٠ ، ١] وتثبيت تغيرات الوزن أثناء تقاربها. [C. A. Collin :2005]

ISSN:2618-0278 Vol. 3 No. 6 June 2021

٣. **الجانب التطبيقي:**

لتقييم أداء نموذجنا المقترح وكذلك لإظهار مزاياه مقارنة بالطرق الأخرى ، يتم إجراء تجارب مختلفة وفقًا لمعايير مختلفة.

3.1 مواصفات مجموعات البيانات المستخدمة يتم تقييم النمـوذج المقترح بثلاث مجموعات بيانات معيارية ، وهي ,ETH-80 [49] :CU3D-100 [17], and ImageNet [50]

تحتوي مجموعة البيانات ETH-80 على ٨٠ كائنًا ثلاثي الأبعاد في ثماني فئات كائنات مختلفة بما في ذلك الموز والسيارة والمواشي والكوب والثعلب والحصان والكمثرى والطماطم. تتضمن كل فئة كائن ١٠ نماذج ويتم تصوير كل نموذج في ٤١ وجهة نظر بزوايا مختلفة. ثلاثة مستويات تصنيف للمجموعة الفرعية المستخدمة من مجموعة البيانات ETH-80 يحتوي المستوى الفائق على فئات الحيوانات والشيء والفواكه. في المستوى الأساسي ، تحتوي فئة الحيوانات على الثعلب والمواشي والحصان ؛ تشمل فئة الكائن الكوب والسيارة ؛ وتشمل فئة الفاكهة الكمثرى والموز والطماطم. يتضمن المستوى الثانوي ثلاثة نماذج لكل فئة على المستوى الأساسي ، أي أن هناك إجمالي ٨ × ٣ = ٢٤ فئة في هذا المستوى.

تم تنظيم CU3D-100 في ١٠٠ فئة بمتوسط ٩ {١٠ نماذج لكل فئة مع تباين متحكم به في الوضع والإضاءة. في هذه الدراسة ، قمنا بالإبلاغ عن النتائج التجريبية على الفئات الست من CU3D-100. ثلاثة مستويات تصنيف للمجموعة الفرعية المستخدمة من مجموعة البيانات CU3D-100. المستوى الأعلى يتضمن فئات المركبات وغير المركبات. في المستوى الأساسي ، تحتوي فئة السيارة على السيارة والطائرة والدراجة النارية ؛ وتشمل فئة غير المركبات مخروط المرور وإشارة المرور وعلامة التحذير. يتضمن المستوى الثانوي أربعة نماذج لكل فئة على المستوى الأساسي ، مما ينتج عنه ٦ × ٤ = ٢٤ فئة في هذا المستوى. ImageNet عبارة عن مجموعة بيانات تضم أكثر من ١٥ مليون صورة عالية الدقة مصنفة تنتمي إلى ما يقرب من ٢٢٠٠٠ فئة [٥٠]. مرة أخرى ، يتم عرض ثلاثة مستويات تصنيف للمجموعة الفرعية المستخدمة من مجموعة بيانات ImageNet. المستوى الأعلى يتضمن فئتي الحيوانات وغير الحيوانية. على المستوى الأساسي ، تحتوي فئة الحيوانات على الطيور والثدييات ؛ وتشمل الفئة غير الحيوانية السيارة والكرة. في المستوى الثانوي ، تشمل فئة الطيور الحمام والبط ؛ تشمل فئة الثدييات القط والكلب ؛ فئة السيارة تشمل الدراجات والدراجات النارية ؛ وتشمل فئة الكرة كرة القدم وكرة التنس ، مما ينتج عنه ٤ × ٢ = ٨ فئات في هذا المستوى.

3.2 **الجانب التجريبي:**

لإجراء التجارب ، يتم تصفية جميع الصور في كل مجموعة بيانات إلى عناصر ذات تردد مكاني منخفض ومتوسط وعالي باستعمال مرشحات Gabor المناسبة. بالنسبة لجميع مجموعات البيانات ، يتم تقسيم صور كل فئة بشكل عشوائي إلى مجموعتين منفصلتين بأحجام متساوية للتدريب والاختبار. لإجراء مقارنة عادلة ، يتم استعمال مجموعات تدريب واختبار متطابقة في جميع النماذج المقارنة. أيضًا ، يتم إجراء التحليل الإحصائي للنتائج التجريبية باستعمال اختبار تحليل التباين (ANOVA).

٣,٣ **نتائج تجريبية**

في التجربة الأولى ، تم التحقق من تأثير تطبيق أنواع مختلفة من الوحدات (RSNN و SNN و CNN و DCNN و HMAX) في كل مستوى تصنيف وتظهر النتائج في الجدول ١. لاحظ أنه في هذه التجربة فقط النماذج التنافسية المختلفة هي فحص في كل مستوى تصنيف. يتم ترشيح الصور المدخلة لهذه الوحدات إلى LSF و ISF و HSF فيما يتعلق بمستوى التصنيف. يتم عرض نتيجة استعمال الوحدة النمطية R-SNN ، والتي تم وصف تفاصيلها في الانموذج المقترح ، في كل مستوى تصنيف في الجدول ١. كما يتم عرض نتيجة استعمال وحدة SNN في كل مستوى تصنيف في هذا الجدول. يتم تدريب SNN المستخدم باستعمال STDP ، بطريقة غير خاضعة للإشراف ، ويمكن حساب ثلاثة أنواع من متجهات الميزات بطول n من طبقة

1 :S_2. متجه السنبلة الأول هو متجه ثنائي ، حيث تكون جميع القيم أصفار ، باستثناء تلك المقابلة للشبكة العصبية ذات السنبلة الأولى .

٢. يخزن متجه عدد السنابل العدد الإجمالي للنبضات المنبعثة من الخلايا العصبية في كل شبكة .

٣. يحتوي المتجه المحتمل على أقصى إمكانات غشاء للخلايا العصبية في كل شبكة. بعد استخراج الميزة لكل من بيانات التدريب والاختبار لكل مستوى تصنيف ، يتم استعمال مصنف SVM لتقييم أداء كل مستوى. بالإضافة إلى ذلك ، في كل مستوى تصنيف ، تتم مقارنة R-SNN بشبكة CNN ضحلة بهيكل مماثل. تحتوي الطبقة الأولى لـ CNN على أربع خرائط ميزات بنفس حجم نافذة الإدخال مثل المرشحات في S_1 لـ R-SNN. بعد ذلك ، يتم استعمال طبقة max-pooling بنفس حجم النافذة وخطوة الخلايا العصبية في C_1 من R-SNN. بعد ذلك ، يتم استعمال طبقة تلافيفية مع خرائط معالم n ويتم تعيين ReLU كوظائف تنشيط للخلايا العصبية التلافيفية. بعد ذلك ، طبقتان كثيفتان: يتم استعمال طبقة مخفية وطبقة ناتجة. في كل مهمة ، يتم ضبط عدد الخلايا العصبية في الطبقة المخفية. تم استعمال وظائف التنشيط ReLU و soft-max للطبقات المخفية والمخرجة ، على التوالي. لتجنب فرط الملاءمة ، يتم إجراء تسوية التسرب على الطبقات الكثيفة ويتم تطبيق تسوية النواة بقيم معلمات مختلفة. أيضًا ، في كل مستوى تصنيف ، تتم مقارنة R-SNN مع Alexnet التي تعد DCNN التي حسنت بشكل كبير دقة التعرف على مجموعة بيانات Imagenet [20]. يتكون Alexnet من خمس طبقات تليها ثلاث طبقات متصلة بالكامل. يتم استعمال مصنف SVM للقيام بالتعرف على الكائن بناءً على متجهات الميزات المستخرجة من الطبقة السابعة من Alexnet المدربة مسبقًا. أيضًا ، تظهر نتائج استعمال وحدة HMAX ، التي تعد أحد النماذج الحسابية الكلاسيكية لعملية التعرف على الكائنات في القشرة البصرية ، في كل مستوى تصنيف في الجدول ١.

كما لوحظ في الجدول ١ ، كان أداء جميع النماذج جيدًا نسبيًا على المستوى الأعلى ، ولكن على المستوى الأساسي وخاصة على مستويات التصنيف الثانوية انخفضت معدلات التعرف عليها. تتفوق الوحدة النمطية R-SNN على SNN بأنواع مختلفة من متجهات الميزات في كل

ISSN:2618-0278 Vol. 3 No. 6 June 2021

مستوى تصنيف. يعد R-SNN خيارًا مناسبًا لتدريب الخلايا العصبية الخاصة بالفئة والتي تكون قادرة على تحديد فئة صورة الإدخال باستعمال زمن الانتقال الأول بدلاً من المصنف الخارجي ، ولكن SNN ليست سوى وحدة استخراج ميزة وهي يتطلب مصنف خارجي. لذلك ، لم يتم فقط زيادة الأداء والمعقولية البيولوجية لـ R-SNN ولكن أيضًا ، انخفضت التكلفة الحسابية. بالإضافة إلى ذلك ، في معظم الحالات ، تعمل R-SNN بشكل أفضل من شبكة CNN الضحلة و Alexnet. إجمالاً ، يبدو أن R-SNN يعد خيارًا جيدًا للاستعمال في وحدات التصنيف ، بناءً على المعقولية البيولوجية وكذلك فعاليتها بالنسبة إلى النماذج الحديثة الأخرى.

الجدول (١) مقارنة النماذج المختلفة بمستويات تصنيف مختلفة في مجموعات البيانات ETH-60 و CU3D و ImageNet

variable	level effect	HMAX	CNN	Alexnet	R-SNN
ETH-60	Superordinate	94.6	91.9	98.3	94.6
ETH-60	Basic	74.3	78.3	64.9	84.7
ETH-60	Subordinate	54.3	44.7	58.3	61.2
CU3D	Superordinate	100	94.7	88.5	94.8
CU3D	Basic	91.4	98.6	97.4	87.6
CU3D	Subordinate	80.3	90.6	80.9	67.3
ImageNet	Superordinate	94	89.32	93.1	89.9
ImageNet	Basic	88.5	86.4	80.23	83.1
ImageNet	Subordinate	54.5	44.9	51.6	65.3

في الانموذج المقترح ، يتم استخدام وحدة R-SNN في كل مستوى تصنيف. للعثور على قيم المعلمات المناسبة ، تم إجراء بحث شامل على كل مستوى تصنيف لمجموعات البيانات المستخدمة.

ومن الجدول ٢ يتم توضيح قيم المعلمات في كل مستوى تصنيف لمجموعات البيانات المذكورة

الجدول (٢) قيم معلمات وحدة R-SNN لكل مستوى تصنيف لمجموعات بيانات ETH-٥٦ و CU3D و ImageNet

		Parameter-R-SNN						
variable	level effect	Weights	Numbers for attributes	$Mr+$	$Mr-$	$Mp+$	$Mp-$	Limit
ETH-60	Superordinate	40	35	0.026	-0.025	0.025	-0.028	160
ETH-60	Basic	33	60	0.0025	-0.023	0.025	-0.025	134
ETH-60	Subordinate	28	214	0.036	-0.005	0.006	-0.006	108
CU3D	Superordinate	40	26	0.012	-0.001	0.002	-0.004	154
CU3D	Basic	35	18	0.0025	-0.032	0.002	-0.052	129
CU3D	Subordinate	24	112	0.021	-0.061	0.006	-0.001	120
ImageNet	Superordinate	36	160	0.041	-0.004	0.021	-0.007	130

ISSN:2618-0278 Vol. 3 No. 6 June 2021

ImageNet	Basic	25	142	0.0024	-0.012	0.0025	-0.052	124
ImageNet	Subordinate	17	123	0.025	-0.034	0.025	-0.004	103

ان تأثيرات كل من التردد المكاني ومستوى التصنيف على أداء الانموذج SNN و R-SNN لكل مجموعة بيانات مذكورة. متوسط دقة التعرف على هذه النماذج (بمتوسط أكثر من ١٠ عمليات تشغيل مستقلة) مع نطاقات كاملة التردد من الصور (أي الصور الأصلية) بالإضافة إلى الصور التي تمت تصفيتها بالتردد للأساسية الفائقة ، ومستويات التصنيف الثانوي ، مقسمة حسب الترددات المكانية المختلفة.

للمستويات السابقة ، مما يعني أن هناك حاجة إلى معلومات تردد أعلى لإجراء التصنيف على المستوى الثانوي. فإن الدقة مع صور النطاق الكامل على من تلك التي تحتوي على نطاقات LSF و ISF و HSF ، لكن اختبارات ANOVA تُظهر أن الاختلافات ليست كبيرة بين النطاق الكامل ونطاقات HSF ($p< 0.05$) للجميع مستويات التصنيف. إجمالاً ، بدلاً من استخدام معلومات النطاقات الكاملة على كل مستوى ، تكون الترددات المكانية المقابلة والمطلوبة لصورة الإدخال فقط كافية

الاستنتاجات

١. تم إهمال بناء نموذج حسابي لتقليد في مستويات التصنيف على أداء الانموذج SNN و R-SNN
٢. تقديم العديد من تجارب علم النفس الفيزيائي في مستويات مختلفة من التجريد واظهرت النتائج تحتوي على نطاقات LSF و ISF و HSF كبيرة بين النطاق الكامل ونطاقات HSF ($p< 0.05$) للجميع مستويات التصنيف.
٣. يحتوي هذا النموذج الهرمي على وحدات فوقية وأساسية وفرعية ، حيث تتكون كل وحدة من وحدة R-SNN لكل من مراحل استخراج الميزات والتصنيف

التوصيات

١. نقترح أنه يمكن استعمال عناصر LSF و ISF و HSF لصورة الإدخال على المستويات الفائقة والأساسية والفرعية ، والتي يمكن أن تقلل بشكل مذهل من الانموذج الحسابي لنموذجنا المقترح.
٢. لتقييم النموذج المقترح يجب تنفيذ تجارب مختلفة على مجموعات البيانات الثلاث المعروفة.
٣. الفصل بين الكائنات الموجودة في المشهد المرئي عن الخلفيات الخاصة بها مثل الصور الموجودة في مجموعة بيانات ImageNet.
٤. يمكن استخدام الانتباه البصري في البحوث المستقبلي ومن المفيد التركيز على الكائنات وميزاتها الإعلامية لتمييز الكائنات في كل مستوى تصنيف.

References

1. A. Yuille, R. Mottaghi, Complexity of representation and inference in compositional models with part sharing, The Journal of Machine Learning Research 17 (1) (2016) 292-319.
2. C. A. Collin, P. A. Mcmullen, Subordinate-level categorization relies on high spatial frequencies to a greater degree than basic-level categorization, Attention, Perception, & Psychophysics 67 (2) (2005) 354-364. doi:10.3758/BF03206498.
3. K. Fukushima, Neocognitron for handwritten digit recognition, Neurocomputing 51 (2003) 161-180. doi:10.1016/S0925-2312(02)00614-8.
4. L. Kauffmann, J. Bourgin, N. Guyader, C. Peyrin, The neural bases of the semantic interference of spatial frequency-based information in scenes, Journal of cognitive neuroscience 27 (12) (2015) 2394-2405. doi:10.1162/ jocn_a_00861.
5. L. Kauffmann, S. Ramano¨el, C. Peyrin, The neural bases of spatial frequency processing during scene perception, Frontiers in integrative neuroscience 8 (2014) 37. doi:10.3389/fnint.2014.00037.
6. M. Craddock, J. Martinovic, M. M. M¨uller, Early and late effects of objecthood and spatial frequency on event-related potentials and gamma band activity, BMC neuroscience 16 (1) (2015) 6. doi:10.1186/ s12868-015-0144-8.
7. S. Grossberg, Towards a unified theory of neocortex: laminar cortical circuits for vision and cognition, Progress in brain research 165 (2007) 79-104. doi:10.1016/S0079-6123(06)650061.
8. T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, T. Poggio, Robust object recognition with cortex-like mechanisms, IEEE transactions on pattern analysis and machine intelligence 29 (3) (2007) 411-426. doi:10.1109/ TPAMI.2007.56.

ISSN:2618-0278 Vol. 3 No. 6 June 2021