تعرّف الوجه باستخدام الشبكات العصبونية وتقدير التقوسات[1]

المهندس محمد أيمن العقاد[2] **الدكتور حسن أبو النور[3]**

الملخص

نعرض في هذه الورقة طريقة لبناء شبكة ثلاثية الأبعاد في منطقة العيون وذلك بتتبع موقع الحدقة وحساب تقوس غاوص Gauss في تلك المنطقة، ثم بالاعتماد على ذلك نرسم الشبكة ثلاثية الأبعاد للوجه البشري. تساق معلومات تقوس غاوص للمناطق المحلية للصورة إلى المصنّف، الذي يعطي شعاع معاملات وزن احتمال انتساب الصورة الأصلية إلى صنف محدد من العدد الكلي للأصناف المتعرف عليها بواسطة النظام.

[1] أعد هذا البحث في سياق رسالة الدكتوراه للمهندس محمد أيمن العقاد بإشراف الأستاذ الدكتور حسن أبو النور.

[2] طالب دكتوراه– قسم الهندسة الالكترونية– كلية الهندسة الميكانيكية و الكهربائية– جامعة دمشق.

[3] أستاذ– قسم الهندسة الالكترونية– كلية الهندسة الميكانيكية و الكهربائية– جامعة دمشق.

1. مقدمة

تندرج معالجة الصور والتعرف، ومنها التعرف على الأشخاص من خلال المميزات الهندسية للوجه، ضمن التطبيقات الرئيسة لأنظمة الشبكات العصبونية وخاصة في:

- استخلاص الخصائص المميزة للصورة وتوصيفها.
- تصنيف الصور أو الخصائص المستخلصة منها.

وتستخدم تقنية الشبكات العصبونية في خوارزميات تحليل الخصائص المميزة للصورة، حيث تتألف الشبكة العصبونية من مجموعة العصبونات المتصلة مع بعضها إذْ إنَّ كل عصبون يستطيع معالجة مجموعة إشارات مطبقة على مداخله ويحولها إلى مجموعة إشارات خرج ذات خصائص محددة. كما تؤدي الروابط بين العصبونات المرمزة بمعاملات الوزن weights الدور الأساسي في عمل هذه الشبكات، حيث تملك الشبكات العصبونية في هذه الحالة إمكانية تعميم جيدة لأنها تستطيع توزيع خبرة التعلم الناتجة عن كل عصبون إلى كامل عناصر الشبكة.

باعتبار أن استخلاص الخصائص الأساسية التي تميز الصورة المرئية يعدُّ أمراً أساسياً فإن تحديد القيم النسبية لهذه الخصائص يتم باختيار معاملات وزن الروابط بين العصبونات والأخذ بالحسبان العلاقات المتبادلة فيما بينها.

تتكون إجراءات التعرف عادة من ثلاث مراحل:

ــ مرحلة قراءة ومعالجة أولية للصورة (رصف الموقع Location alignment، تجزيء Segmentation، إزالة الضجيج، وتحسينات إضافية أخرى).

ــ مرحلة توصيف الصورة (استخلاص الخصائص المميزة للصورة في ضوء نماذج مخزنة، مثلاً نماذج لانحناء السطوح).

ــ مرحلة تصنيف باستخدام خوارزميات الشبكات العصبونية (حيث يتم تحليل الخصائص المستخلصة واتخاذ قرار التعرف النهائي).

1.1. استعراض طرق الشبكات العصبونية

في البداية اعتمدت طرق الشبكات العصبونية على خارطة كوهنين Kohonen المشتركة associative map [3]، واستخدمت عدداً قليلاً من صور الوجوه، وتم من خلالها الإعلان عن تعرف دقيق حتى بحالة صور دخل مضجة ومشوهة. ولقد تم أيضاً إظهار هذه الإمكانية باستخدام العتاد الضوئي optical hardware [4].

لاحقاً لذلك تم استخدام الشبكات العصبونية ضمن طريقة من أجل تصنيف الأشخاص وفقاً لجنسهم gender [5]، والتي استخرجت آلياً شعاع خصائص ذا 16 بعداً مثل ثخانة الحاجبين، عرض الأنف والفم، وستة معايير للذقن chin radii، الخ. تم تدريب شبكتي تابع فوقي Hyper Basis Function (HyperBF) [6] واحدة لكل جنس. تم تقييس صور الدخل بالنسبة للتدريج scale والدوران باستخدام مواقع العيون، والتي حددت آلياً. بعد مقارنة مخارج الشبكتين HyperBF، تحدد الشبكة ذات الخرج الأكبر دليل الجنس gender label في الصورة المختبرة. تم في تجارب التصنيف الفعلية استخدام مجموعة جزئية فقط من شعاع الخصائص ذي ال 16 بعداً. تألفت قاعدة المعطيات من 21 ذكراً و 21 أنثى. من أجل التصنيف تم استخدام استراتيجية "ترك واحد في الخارج" leave-one-out [2]. عندما استخدم شعاع الخصائص من مجموعة التدريب كشعاع اختبار، تم الإعلان عن دقة تعرف صحيح قيمتها 92.5%؛ وانخفضت الدقة إلى 87.5% من أجل وجوه ليست في مجموعة التدريب. تم إثبات صحة validation بعض نتائج التصنيف الآلي من خلال استخدام التعرف الإنساني using humans.

تم توسيع طريقة التصنيف وفقاً للجنس إلى تعرف الوجه باستخدام شعاع خصائص موسع ذي 35 بعداً و شبكة HyperBF لكل شخص. المسوغ للبنية التحتية underlying structure هو مفهوم العصبون الجد grandmother neuron: وهو

عصبون مفرد (تابع غاوص في الشبكة HyperBF) لكل شخص. ولوجود صور تدريب قليلة نسبياً لكل شخص، تم توليد قاعدة معطيات تركيبية synthetic database بواسطة التأثير في متوسط أشعة الخصائص لصور الأشخاص المتوافرة وهذه الصور استخدمت كعينات اختبار.

تم الحصول على نتائج التصنيف من أجل مجموعات مختلفة من معاملات التوليف (مكافئات coefficients، مراكز centers ومقاييس metrics ال HyperBF). تم أيضاً تقديم بعض المشاركة في شرح التصرف الكاريكاتيري caricatural behavior لشبكات HyperBF من قبل الدراسات النفس بدنية psychophysical.

بنيت الطرق [7] و [8] على بنية الرابطة الديناميكية Dynamic Link Architecture (DLA)، في محاولة لحل المشكلة الأساسية لتعبير العلاقات السينتاكتيكية[1] syntactical في الشبكات العصبونية التقليدية. استخدمت في هاتين الطريقتين مويجات جيبر Gabor wavelets للخصائص. تستخدم بنى الروابط الديناميكية الليونة السينابتيكية[2] synaptic وهي قادرة فوراً على تشكيل مجموعات عصبونات مجمعة ضمن مخططات بنيوية والحفاظ على ميزات الأنظمة العصبونية. من الضروري للطرق التي تستخدم بنى الروابط الديناميكية توافر حيز صور وحيز نماذج. يماثل حيز الصور مناطق الرؤية اللحائية و حيز النماذج يماثل اللحاء الزمني البيني في نظام الرؤية الأحيائي. تم مؤخراً توسيع بنى الروابط الديناميكية إلى ما يدعى مطابقة

[1] سينتاكتيكي: عائد إلى قواعد تركيب الأشياء في أشكالها الصحيحة ضمن إطارها العام.

[2] سينابتيكي: عائد إلى تغير قوة الإشارة عبر نقاط الاشتباك العصبي التي تربط الخلايا العصبية ببعضها.

المخطط لين الحزمة Elastic Bunch Graph Matching (EBGM) [9،10]، وهي مماثلة للطريقة الموصوفة أعلاه ولكن بربط مجموعة أنابيب jets لكل عقدة بدلاً من واحد، كل أنبوب مستنبط من صورة وجهية مختلفة. للتصدي لمشكلة تغير الوضعية pose variation، يتم أولاً تحديد وضعية الوجه باستخدام معلومات مسبقة [1] ويتم تعلم تحولات transformations المجموعات sets بالنسبة لتغيرات الوضعية [11].

2. التوصيف

إن عملية التوصيف يجب أن تتضمن أكبر عدد ممكن من الخصائص المميزة لصورة وجه شخص عن الآخر، و تعتمد الخوارزمية العامة للتوصيف عند بناء شبكة كثيرة الأضلاع في منطقة العيون وفق الشكل 1 على العمليات الآتية:

آ ــ اكتشاف حدقتي العينين ورسم الشبكة بالنسبة لهما.

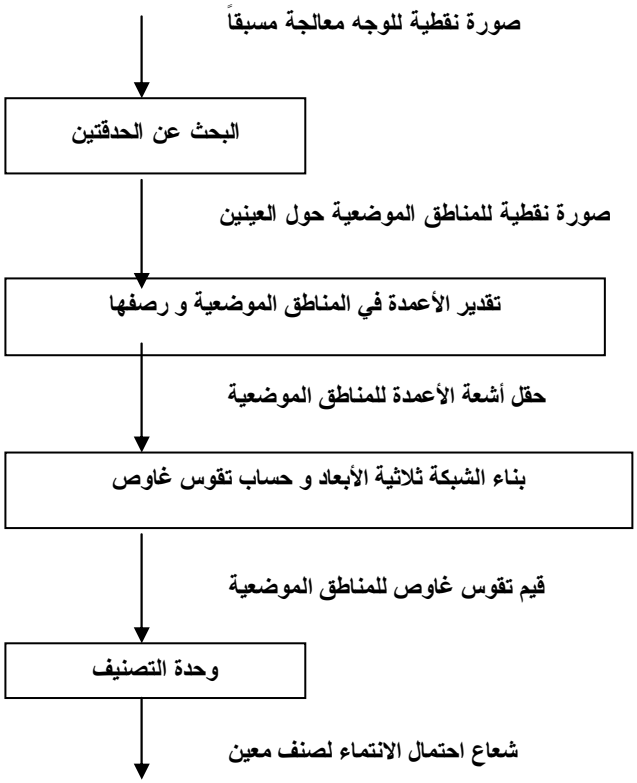
ب ــ حساب تقوس غاوص في هذه المنطقة.

معطيات الدخل لهذه الشبكة: خرج وحدة القراءة والمعالجة الأولية.

معطيات الخرج لهذه الشبكة: قيم تقوس غاوص للمناطق الموضعية local regions حول العينين.

إن مصفوفة العين اليسرى بحجم 7×5 وعناصرها $L_{ij} \in R$ حيث

$i \in [1, 5]$ و $j \in [1, 7]$.



الشكل 1: خوارزمية استخلاص خصائص الصورة.

1.2. البحث عن الحدقتين (تمركز العينين)

إن العينين هما العنصران الأساسيان في الوجه اللذان يسهمان في التعرف الآلي على الوجه وذلك لأن المسافة بين الحدقتين لا تتغير مع تغير تعابير الوجه في حال النظـــر

إلى مسافة بعيدة نسبياً. ومن ثَمَّ فيمكن عملياً استخدام هذه المسافة لتحديد بقية مواقع عناصر الوجه الأخرى.

إن متطلبات قياس المسافة بين الحدقتين تفرض دقة قياس عالية لموقع الحدقتين. لهذا السبب تم بناء خوارزمية خاصة تعتمد على إيجاد زوج من النهايات الصغرى المحلية مع إمكانية استخدام طرائق مقارنة النماذج.

معطيات الدخل: صورة نقطية معالجة.

مصفوفة الصورة النقطية المعالجة ذات حجم W x H، عناصرها $E_{ij} \in Z^+$ حيث $I \in [1, H]$ و $j \in [1, W]$ و $W,H \in Z^+$ $W=H=200$.

معطيات الخرج: صور نقطية للمناطق الموضعية حول العينين. مصفوفة العين اليسرى بحجم 70×50 وعناصرها $L_{ij} \in Z^+$ حيث $i \in [1, 50]$ و $j \in [1, 70]$.

يبين الشكل 2 صورة أصلية، الصورة النقطية بمصفوفة 200×200 المولدة من النظام، و استخدام نافذة 70×50 حول منطقة العين اليسرى.

الخوارزمية المنفذة:

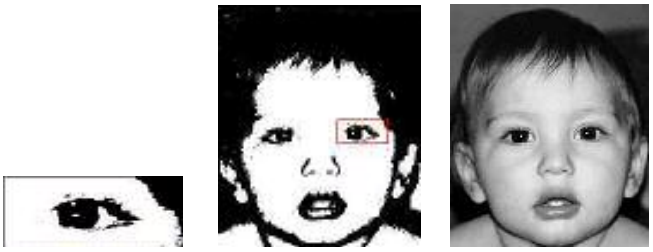
تتضمن هذه الخوارزمية مرحلتين:

آ. ترشيح العتبة: يبنى تحديد موقع الحدقتين على فرضية أن الحدقة عموماً هي قرص معتم مع حافات محيطة مضيئة. تجزأ صورة الوجه بعتبة منخفضة جداً لعزل 5% من النقاط الأكثر تعتيماً [13,14]. يمكن أن تتبع هذه النقاط إلى الحدقتين كما هو مبين في الشكل2.

تحسب العتبة كما يأتي: $T = \dfrac{\sum_{i=1}^{H}\sum_{j=1}^{W} E_{ij}}{H * W * 3}$ حيث E_{ij} هو سطوع النقطة ذات الموضع (i,j) ، ثم ننشئ صورة نقطية مرشحة ثنائية التدرج:

إذا كانت $E_{ij} < T$ فإن $E_{ij} = 0$

وإلا $E_{ij} = 1$



(ج) (ب) (آ)

الشكل 2: آ- **صورة أصلية من قاعدة المعطيات، ب-الصورة النقطية 200×200 المولدة من النظام، و النافذة 50×70 حول منطقة العين اليسرى، و ج- قرص الحدقة المرشح للعين اليسرى.**

ب. مقارنة النموذج: تتم مقارنة النموذج template مع العين اليسرى ويكرر الإجراء نفسه بالنسبة للعين اليمنى.

إن الترابط Correlation بين نقاط D_{ij} ونقاط الصورة المرشحة F_{ij} هو عدد النقـــاط المتطابقة الذي يحسب كما يأتي [20]:

$$C= \sum_{i,j\in w} F_{ij}\Big|_{F_{ij}=D_{ij}=0}$$

حيث w النافذة اليمنى أو اليسرى للحدقتين

وحيث D_{ij} توصف كما يأتي:

$$\text{إذا} \quad (i-x_0)^2+(j-y_0)^2 < r^2 \qquad D_{ij=0}$$

$$\text{وإلا} \qquad D_{ij=1}$$

ونظراً لأنَّ الدائرة ذات المركز (x_0,y_0) ونصف القطر r توصف كما يأتي:

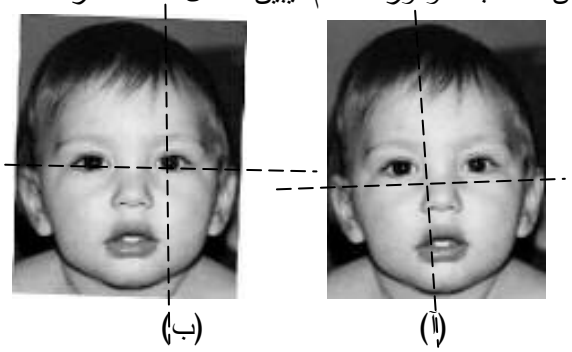
للعين اليمنى $x_0 \in (W/2,W-(20+r))$ للعين اليسرى $x_0 \in (20+r,W/2)$

و $y_0 \in (20+r,H/2-r)$ و $r \in [8,12]$

من قيم x_0, y_0 و r بالإضافة إلى القيمة العظمى للترابط C_{max} نتوصل إلى تحديد مركزي الحدقتين الافتراضيين x_h, y_h وتكون المصفوفة الناتجة $L_{i,j} = E_{i,j}$ حيث $i \in [x_h-25, x_h+25]$ و $j \in [y_h-35, y_h+35]$.

2.2. رصف صورة الوجه:

نرسم في البداية خطاً بين مركزي العينين (النقطتين البيض داخل البؤبؤين) ندعوه خط مركزي العينين و نرسم خطاً عامودياً عليه يمر عبر رأس الأنف ندعوه خط عامود الوجه. من ثم نرصف align صورة الوجه وفقاً لخط مركزي العينين بتدوير نافذة الوجه بمقدار الزاوية بين الخط الأفقي و خط مركزي العينين، بحيث يصبح خط مركزي العينين متطابقاً مع المحور x. يوحد هذا الرصف انحناء الوجه بين جميع الصور ويخفض الحسابات ودورة التعلم. يبين الشكل 3 مثالاً لرصف صورة.



الشكل 3: (آ) صورة عليها خطوط الرصف الرئيسية، (ب) الصورة المرصوفة بالدوران بمقدار 2 درجة.

يمكن الاعتماد على النقطتين البيض داخل البؤبؤين في كثير من الأحوال وخصوصاً في حال الإضاءة المباشرة، وللحصول على عمومية أكثر يمكن تحديد مركزي القرصين السوداويين الممثلين للبؤبؤين حسابياً.

3.2. رسم الشبكة كثيرة الأضلاع:

يتم من أجل رسم الشبكة كثيرة الأضلاع تحقيق ما يلي:

– حساب أشعة الأعمدة normals على السطوح في صورة الوجه باستخدام طريقة هورن Horn التكرارية [12,13,14].

– حساب عمق الشبكة Z بناءً على أشعة الأعمدة.

1.3.2. التحديد التكراري للأعمدة (إيجاد أشعة الأعمدة ورصفها)

تناسب هذه الطريقة السطوح الناثرة للضوء بشكل انتشاري والتي لها قابلية العكس نفسها أو الامتصاص (albedo) على كامل المساحة (حسب لامبرت Lambert) وبجهة منبع إضاءة معروفة. كما أن إعادة بناء الشكل من الصورة ستتم من أجل صور مضجة (حقيقية).

معطيات الدخل: صور نقطية للمناطق الموضعية حول العينين.

مصفوفة العين اليسرى بحجم 50 x 70 وعناصرها $L_{ij} \in Z^+$ حيث $i \in [1, 50]$ و $j \in [1, 70]$.

معطيات الخرج: حقل أشعة الأعمدة للمناطق الموضعية حول العينين.

المصفوفة للعين اليسرى بحجم 50 x 70 وعناصرها $N_{ij} \in L$ حيث $[1, 50]$ $i \in$ و $[1, 70]$ $j \in$ و L هي الفضاء ثلاثي الأبعاد.

الخوارزمية: نستخدم في هذه الطريقة نظام معادلات تكرارياً لحساب أشعة الأعمدة في كل نقطة من الصورة باستخدام قيم سطوع معروفة:

$$m_{ij}^{k+1} = \bar{n}_{ij}^{\,k} + e^2 (E_{ij} - n_{ij}^k . s)\, s / 4\lambda$$

$$n_{ij}^{k+1} = m_{ij}^{k+1} / |m_{ij}^{k+1}|$$

حيث m_{ij}^{k+1} شعاع العامود في النقطة i,j عند التكرار $k+1$

n_{ij}^{k} شعاع العامود المقيس في النقطة i,j عند التكرار k

n_{ij}^{k+1} شعاع العامود المقيس في النقطة i,j عند التكرار $k+1$

S شعاع منبع الضوء ويساوي $(1، 0.001، 0.001)^T$

E_{ij} شدة سطوع الصورة في النقطة i,j

e المسافة بين خلايا متجاورة في الصورة

$\ddot{e}$ معامل طاقة تعديل العامود

$\bar{n}_{ij}$ العامود الوسطي المحلي في النقطة i,j :

$$\bar{n}_{ij} = \frac{1}{4}(n_{i,j+1} + n_{i,j-1} + n_{i+1,j} + n_{i-1,j})$$

نحسب أشعة الأعمدة لكل نقطة من صورة العين اليسرى حيث القيمة الابتدائية لأشعة الأعمدة $(0,0,1)$. وقد أثبتت التجارب بالنسبة لأغراضنا أن *أربعة* تكرارات كافية أي $k \in [0,3]$.

بعد استكمال دورات التكرار نقوم برصف الأعمدة، ونوجد العامود الوسطي كما يأتي:

$n^{(3)}_{i/2,j/2} = (0,0,1)^T$ حيث $n_{avg} = n^{(3)}_{i/2,j/2}$

كما أن أي عامود في الصورة سيكون: $N_{i,j} = n^{(3)}_{i,j} - n_{avg}$

2.3.2. تحديد تقوس السطح (بناء الشبكة ثلاثية الأبعاد، وحساب الزوايا والتقوس)

إذا حدد السطح بالمعادلة $z=f(x,y)$ فإن R_1 و R_2 هما نصفا القطرين الأساسيين للتقوس ويحسبان كجذور لمعادلات تربيعية [21]:

$$(rt - s^2) R^2 + h [2pqs - (1+p^2) t - (1+q^2) r] R + h^2 = 0$$

حيث $p=dz/dx,\ q=dz/dy,\ r=d^2z/dx^2,\ s=d^2z/dxdy,\ t=d^2z/d^2y$

و $h = \sqrt{1 + p^2 + q^2}$

وباعتبار أن التقوس الوسطي للسطح في نقطة هو $H = ½ (1/R_1 + 1/R_2)$

وأن تقوس غاوص هو $K = 1/R_1R_2$

(من أجل اسطوانة بنصف قطر a يكون $K = \frac{1}{2a}$ و أن $K > 0$ من أجل نقاط قطع ناقص ومن أجل نقاط قطع زائد $K < 0$ و من أجل نقاط قطع مكافئ $K = 0$)

فإذا حدد السطح بالمعادلة $z=f(x,y)$ فإن H و K يحسبان وفقاً للمعادلات الآتية:

$$H = r(1 + q^2) - 2\,p\,q\,s + t\,(1 + p^2) / 2(1 + p^2 + q^2)^{3/2}$$
$$K = r\,t - s^2 / (1 + p^2 + q^2)^2$$

المطلوب الآن حساب المشتقات p, q, r, s, t ونقترح لذلك استخدام تقريب منفصل discrete approximation للسطح باستخدام طريقة الفروق النهائية final differences [22]، ومن ثَمَّ تحويل تابع عمق الشبكة Z net depth function إلى تابع سطح مستمر z . تحدد Z بشكل منفصل على كامل مجال السطح بالخطوة الواحدية للشبكة unit net step. كما نقترح استخدام تقريب منفصل من أجل p و q :

$$P = dz/dx = Z_{x,y} - Z_{x-1,y}$$
$$q = dz/dy = Z_{x,y} - Z_{x,y-1}$$

بدلاً من مشتقات يسارية الفرق left difference derivatives نستخدم مشتقات مركزية الفرق central difference derivatives ونحصل في أثناء التقريب الخطي للسطح الموصوف بخارطة الأعماق Z على نموذج من النوع المتعارض cross type:

$$P = dz/dx = Z_{x+1,y} - Z_{x-1,y}/2$$
$$q = dz/dy = Z_{x,y+1} - Z_{x,y-1}/2$$

الآن وباستخدام قيم المشتقات الجزئية الأولى سنستنتج معادلات حساب المشتقات الجزئية الثانية

$$r = d^2z/dx^2 = p_{x+1,y} - p_{x-1,y} / 2 = Z_{x+2,y} - Z_{x,y} / 4 - Z_{x,y} - Z_{x-2,y} / 4$$
$$s = d^2z/dxdy = p_{x+1,y} - p_{x,y-1} / 2 = Z_{x+2,y} - Z_{x,y} / 4 - Z_{x,y} - Z_{x,y-2} / 4$$
$$t = d^2z/d^2y = p_{x,y+1} - p_{x,y-1} / 2 = Z_{x,y+2} - Z_{x,y} / 4 - Z_{x,y} - Z_{x,y-2} / 4$$

وبحساب قيم p, q, r, s, t من الممكن حساب H و K في كل نقطة من نقاط الصورة.

معطيات الدخل: حقل أشعة الأعمدة للمناطق الموضعية حول العينين.

المصفوفة للعين اليسرى بحجم 70×50 وعناصرها $N_{ij} \in L$ حيث $[1, 50]$ $\ni i$ و $[1, 70] \ni j$ و L هي الفضاء ثلاثي الأبعاد.

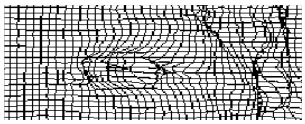
معطيات الخرج: قيم تقوس غاوص للمناطق الموضعية حول العينين.

المصفوفة للعين اليسرى بحجم 7×5 وعناصرها $K_{ij} \in R$ حيث $[1, 5] \ni i$ و $[1, 7] \ni j$.

الخوارزمية: تبنى الشبكة ثلاثية الأبعاد (خارطة الأعماق) بحساب الإحداثي z في كل نقطة

$$N_{ij} = \{LEN_x, LEN_y, LEN_z\} \; . \; Z_{ij} \in R$$
$$z_0 = LEN_{z11}$$
$$Z_{ij} = LEN_{zij} + z_0$$
$$z_0 = LEN_{zij}$$

كما هو مبين في الشكل 4 الشبكة ثلاثية الأبعاد للمناطق المحلية حول العين اليسرى.



الشكل 4: الشبكة ثلاثية الأبعاد للعين اليسرى.

3. وحدة المصنف

معطيات الدخل: قيم تقوس غاوص للمناطق الموضعية حول العينين.

المصفوفة للعين اليسرى بحجم 7×5 وعناصرها $K_{ij} \in R$ حيث $[1, 5] \ni i$ و $[1, 7] \ni j$.

معطيات الخرج:

شعاع صف معاملات الوزن لاحتمال انتماء الصورة الأصلية لصنف معين ذي بعد $N \in Class\ Amount$، من العدد الكلي للأصناف المتعرف عليها بواسطة النظام.

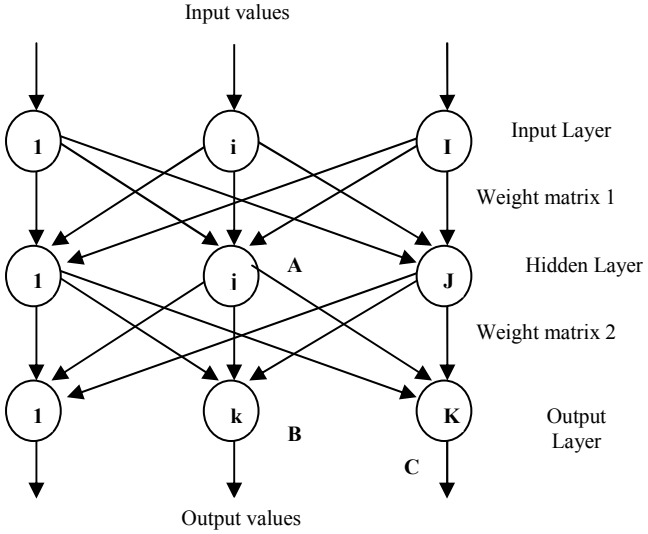
$Output_{ca} \in R$ حيث $ca \in [1, Class\ Amount]$

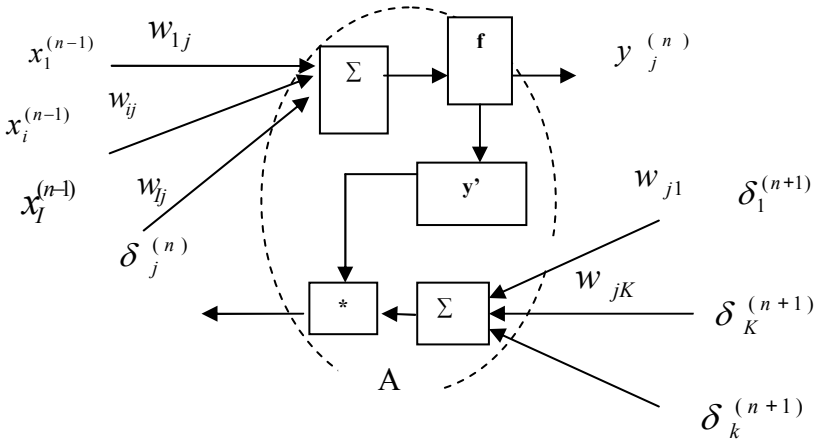
تذهب هنا معطيات تقوس غاوص للمناطق المحلية للصورة إلى المصنّف. المصنّف هو نموذج حسابي يستخدم الشبكات العصبونية الصنعية المؤلفة من عناصر بسيطة (الخلايا العصبية) وروابط بين مخارجها (نقاط الاشتباك العصبي). ونظراً لأنَّ الشبكات العصبونية الصنعية مستقرة تجاه تحويل الإحداثيات والضجيج فقد تم تحقيق المصنّف على أساس توزع البرسبترون perceptron متعددة الطبقات [15,16,17]. المداخل عددها $k_0 = 35$ و الخلايا العصبية $k_1 = 70$ ضمن طبقة مخفية واحدة ويتطابق عدد الخلايا العصبية في طبقة الخرج مع عدد الأصناف المطلوب التعرف عليها $k_2 = Class\ Amount$.

إن التعلم يتم باستخدام طريقة التوزيع العكسي للخطأ reverse error distribution [23]. وكتابع تخفيض reducing function (رمزي المنطق logistic) يؤخذ منحنى بشكل حرف C (sigmoid):

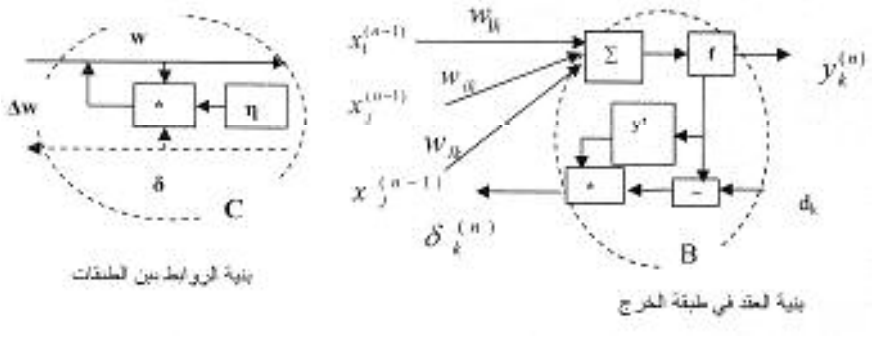
$$y_j^{(n)} = \frac{1}{1 + e^{s_j^{(n)}}}$$

إن بنية البرسبترون متعدد الطبقات و مخطط الجريان ضمن الشبكة في أثناء التعلم باستخدام طريقة التوزع العكسي للخطأ مبينة في الشكل 5.





بنية العقد في الطبقة المخفية



الشكل 5: بنية البرسبترون متعدد الطبقات و مخطط الجريان ضمن الشبكة أثناء التعلم باستخدام طريقة التوزع العكسي للخطأ.

1.3. خوارزمية تعليم الشبكة العصبونية: بالاستعانة بإجرائية التوزيع العكسي يتم بناء الخوارزمية كما يأتي:

أ. نرسل الصورة إلى مداخل الشبكة ونحسب قيم المخارج:

$s_j^{(n)} = \sum_{i=0}^{M} y_i^{(n-1)} . w_{ij}^{(n)}$ و $y_j^{(n)}$ حيث M عدد الخلايا العصبونية في $n\text{-}1$ طبقة.

$y_j^{(n-1)} = x_{ij}^{(n)}$ العصبون رقم i المدخل j للطبقة n

$y_j^{(n)} = f(s_j^{(n)})$ حيث التابع f منحنى بشكل حرف C للقيمة 5.29

$y_j^{(0)} = I_q$ حيث I_q عنصر شعاع صورة الدخل رقم q

$y_q^{(0)} = K_{m,n}$ للعين اليسرى و $q \in [1,35]$

حيث $n = 1, 2$ و j كمية العصبونات في الطبقة n: $j = 35, 70, Class\ Amount$

ب. حساب $\delta^{(N)}$ و $\delta^{(2)}$ لطبقة الخرج:

$$\delta_l^{(N)} = (y_i^{(N)} - d_l).\frac{dy_l}{ds_l}$$

$$l \in [1,k_2] , \quad \frac{dy_l}{ds_l} = dy_l(1-dy_l) \quad \text{حيث} \quad \delta_l^{(2)} = (y_i^{(2)} - d_l).y_l(1-y_l)$$

إذا كان العصبون على الخرج يتطابق مع صنف معين فإنّ $dl = 1$

إذا كان العصبون على الخرج لا يتطابق مع صنف معين فإنّ $dl = 0$

لحساب الأوزان و$w^{(N)}$ لتغير N طبقة: $\Delta w_{ij}^{(N)} = -\eta.\delta_j^{(N)}.y_i^{N-1}$

لحساب الأوزان و$w^{(2)}$ لتغير 2 طبقة: $\Delta w_{ij}^{(2)} = -\eta.\delta_j^{(2)}.y_i^{(1)}$

حيث $i \in [1,k_2] , j \in [1,k_1]$

ت. حساب الأوزان $\delta(n)$ و $\Delta w(2)$ من أجل كل الطبقات المتبقية:

$$\delta_j^{(n)} = [\sum_k \delta_k^{(n+1)}.w_{jk}^{(n+1)}].\frac{dy_j}{ds_j}$$

$$dy_j/ds_j = dy_j(1-dy_j)$$

$$\delta_j^{(n)} = [\sum_k \delta_k^{(n+1)}.w_{jk}^{(n+1)}].y_j(1-y_j)$$

$$\Delta w_{ij}^{(n)} = -\eta.\delta_j^{(n)}.y_i^{(n-1)}$$

ث. تصحيح كل الأوزان في الشبكة العصبونية:

$$w_{ij}^{(n)}(t) = w_{ij}^{(n)}(t-1) + \Delta w_{ij}^{(n)}(t)$$

حيث t دورة التعلم époque $t \in [1,500]$

تنفذ في عملية التعرف الخطوة آ وتعطى النتيجة: $Output_{ca} = y_{ca}^{(2)}$

إذا كان الخطأ جوهرياً فيتم تكرار الذهاب إلى الخطوة آ وإلا يتم إنهاء الحلقة. إن الشبكة تمر بالخطوة آ على كل الصور التدريبية، والخرج هو رقم الانتساب للنوع حيث النتيجة تعطى كما ما يأتي:

$A \in$ [عدد الأنواع, 1] .

2.3. تقارب الشبكة العصبونية:

تتعلق بالتعلم مسألة مهمة هي مشكلة تقارب الخطأ أي تخفيض الخطأ بين قيم الوحدات المرغوب فيها والمحسوبة منها. الهدف هو تحديد مجموعة الأوزان التي تخفض الخطأ.

شبكات البرسبترون العصبونية متعدد الطبقات معروفة بامتلاكها معدل تقارب أبطأ من المرغوب فيه ومما يقود إلى أزمان تدريب طويلة.

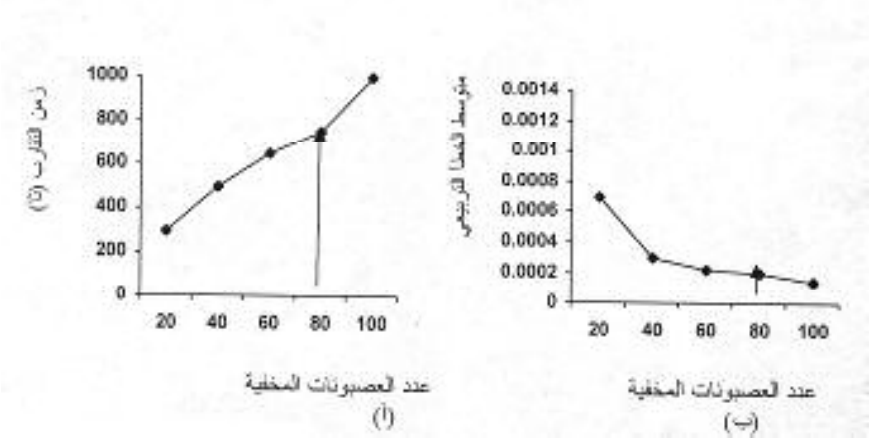
يتبع فشل التقارب لبنية شبكة عصبونية غير صحيحة مع طبقات أوزان إما قليلة جداً أو كثيرة جداً، أو وجود عصبونات مخفية قليلة جداً. و عادةً لا يتم التقارب بسبب تشبع بعض العصبونات قبل غيرها. يحدث التشبع إذا دفع الخرج نحو قيمته العظمى عند نقطة معينة قبل الوصول إلى التقارب، ومن ثَمَّ المشتق صغير جداً من أجل القيام بتغيرات وزن ذات قيمة، دافعاً الشبكة للاستقرار عند قيمة محلية صغرى غير صحيحة أو بلوغ حالة من الشلل. تتحدد تغيرات الوزن من أجل كل عصبون بواسطة ضرب معدل التعلم (ثابت لكامل الشبكة) بحد الخطأ (ثابت لكل عصبون عند أي عبور) و المداخل المعنية [24,25].

ولدراسة تقارب الشبكة وسلوكها في أثناء التدريب وكفاءة التعميم فيها، نقوم بتغيير المحددات بشكل محكوم بحيث نغير أحد المحددات ونثبت البقية كل على حدة، ونحسب الخطأ، من ثم نختار قيم المحددات التي تعطي قيمة خطأ و زمن تقارب منخفضين نسبياً وفقاً لمبدأ المقايضة بين الاثنين. وبعد دراسة التقارب آخذين تجنب حدوث الإشباع بالحسبان حصلنا على أفضل النتائج بعد أن قمنا بما يأتي:

1. خفضنا مجال الوزن الابتدائي عندما ازداد عدد مداخل العصبون.
2. استخدمنا عدداً كافياً من الخانات الثنائية من أجل التغيرات في الوزن.
3. أبقينا معدل التعلم أصغر ما يمكن.

4. زدنا بشكل وسطي عدد المرات التي يتم فيها تمثيل نماذج محددة لكل دور وبحيث يتم تمثيل كل الأصناف بالشكل المناسب.

5. توقفنا عن التدريب لدى بلوغ جميع مخارج الشبكة حالتها النهائية المطلوبة لجميع النماذج ضمن تسامح محدد.



الشكل 7: آ– أزمان التقارب لنظام تعرف الوجه، ب – متوسط الخطأ التربيعي نسبةً لعدد العقد المخفية.

حيث لاحظنا بالنسبة لأغراضنا أن أفضل قيمة لعدد العصبونات المخفية هي 70 عند عدد مداخل قيمته 35.

3.3. قاعدة المعطيات:

إن قاعدة معطيات تدريب الشبكة العصبونية عبارة عن قاعدة معطيات غرضية التوجه تتألف بالأساس من 216 صورة، حيث توجد صورة واحدة لكل شخص في أثناء التعرف. إن صور الوجوه عبارة عن أصناف أساسية والتي تتألف من أصناف فرعية والتي هي فعلياً عناصر الوجه. يتم توليد عدة صور لكل شخص بواسطة عملية تركيب من الصور الأساسية [13] في أثناء التدريب وبنمط واحد – لأكثر.

4. النتائج التجريبية:

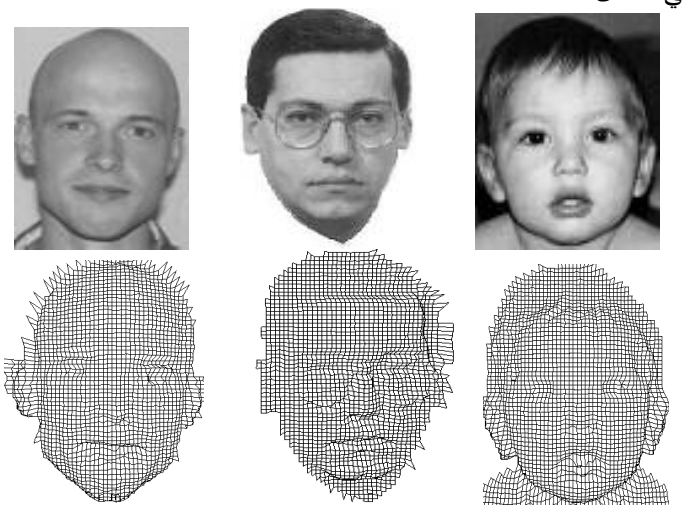
في البداية لاختبار نظامنا أجرينا التجربة الآتية:

تضمنت مجموعة صور الاختبار 15 صورة وجه، عولجت معالجة أولية وخفضت إلى قياس 200x200 بيكسل و 16 مليون تدرج رماديات، العينات موضحة على الشكل 6. الصور في شروط أخذ مختلفة، زوايا استدارة وجه مختلفة، شروط إضاءة، وتعابير وجه.

في البداية في بعض العينات، لم يعثر على الحدقتين، وفي عينات أخرى لم يكن تحديد الحدقتين صحيحاً تماماً، لعل ذلك بسبب الإضاءة غير المنتظمة، مما قادنا إلى تسوية النصوع في عينات الصور وتدريج الصور إلى القياس نفسه.

بعد عملية المعالجة الأولية لاحظنا تحسيناً ملحوظاً في تحديد الحدقتين ومن ثَمَّ في الأداء الكلي للنظام.

بينت النتائج أنَّ نظامنا واعد، ويقود إلى نتائج مرضية في التعرف على الوجه كما هو موضح في الشكل 6.



الشكل 6: بعض العينات من الشبكات ثلاثية الأبعاد المبنية بواسطة النظام.

5. خاتمة

بينا طريقة للتعرف على الأشخاص من صورة الوجه على أساس الشبكات العصبونية حيث تم تناول ما يأتي:

- خوارزمية البحث عن الحدقتين.
- تعيين التقوس التوسيطي و الغاوصي.
- تصنيف الصور على أساس توزع الشبكات العصبونية متعددة الطبقات (البرسبترون).

تتميز الشبكات العصبونية بآليات التعلم الشمولية وقدرة كل عناصرها بالعمل على التفرع مما يزيد الكفاءة في حل المهمة بالمقارنة مع الطرائق الأخرى مثل الطرائق الاحتمالية، الفرز الخطي، شجرات الحسم، الخ.

1.5. مقارنة مع الطرائق الأخرى:

عادة طرائقً NN مخصصة إما لتطبيقات الزمن الحقيقي [19] أو تطبيقات في غير الزمن الحقيقي [18]. تستخدم أيضاُ طرائق ال NN أكثر من صورة في قاعدة المعطيات لكل شخص [15]. طريقتنا جيدة لكل من تطبيقات الزمن الحقيقي و غير الزمن الحقيقي، ولا تحتاج إلى أكثر من صورة واحدة لكل شخص. هذا يستهلك حيز تخزين أقل في نظام قاعدة المعطيات، ويجعل طريقتنا أسرع وأكثر مرونة.

المراجع

[1] N. Kruger, M. Potzsch, and C.v.d. Malsburg, "Determination of Face Position and Pose with a Learned Representation Based on Labelled Graphs", Image and Vision Computing, Vol. 15, pp. 665-673, 1997.
[2] K. Fukunaga, Statistical Pattern Recognition, New York: Academic Press, 1989.
[3] T. Kohonen, Self-Organization and Associative Memory, Berlin: Springer-Verlag, 1988.
[4] Y. S. Abu-Mostafa and D. Psaltis, "Optical Neural Computers", Scientific American, Vol. 256, pp. 88-95, 1987.
[5] R. Brunelli and T. Poggio, "HyperBF Networks for Gender Classification", in Proceedings, DARPA Image Understanding Workshop, pp. 311-314, 1992.
[6] T. Poggio and F. Girosi, "Networks for Approximation and Learning", Proc. IEEE, Vol. 78, pp. 1481-1497, 1990.
[7] J. Buhmann, M. Lades, and C.v.d. Malsburg, "Size and Distortion Invariant Object Recognition by Hierarchical Graph Matching", in Proceedings, International Joint Conference on Neural Networks, pp. 411-416, 1990.
[8] M. Ladesraj, J. Vorbruggen, J. Buhmann, J. Lange, C.v.d. Malsburg, and R. Wurtz, "Distortion Invariant Object Recognition in the Dynamic Link Architecture", IEEE Trans, on Computers, Vol. 42, pp. 300-311, 1993.
[9] L. Wiskott, J.M. Fellous, N. Kruger, and C.v.d. Malsburg, "Face Recognition and Gender Determination", in Proceedings, International Workshop on Automatic Face and Gesture Recognition, pp. 92-97, 1995.
[10] L. Wiskott, J.M. Fellous, and C.v.d. Malsburg, "Face Recognition by Elastic Bunch Graph Matching", IEEE Trans, on Pattern Analysis and Machine Intelligence, Vol. 19, pp. 775-779, 1997.
[11] T. Maurer and C.v.d. Malsburg, "Single-View Based Recognition of Faces Rotated in Depth", in Proceedings, International Workshop on Automatic Face and Gesture Recognition, pp. 176-181, 1996.
[12] M. J. Brooks and B. K. P. Horn, Shape and Source from Shading. Number AIM-820. Artificial Intelligence Lab., MIT press, 1985.
[13] M. A. Al-Akkad and V. N. Kochuhanov, “A framework for building a general 3D model”, Izhevsk, May 2003.
[14] G.G. Gordon, “Face recognition based on depth and curvature features”, In Proc. IEEE Conf. Comput. Vis. Patt. Recogn., pp. 808–809. 1992.
[15] A. C. Tsoi, S. Lawrence, C. L. Giles and A. D. Back, “Face Recognition: A Convolutional Neural Network Approach”, IEEE Transactions on Neural Networks, Special Issue on Neural Networks and Pattern Recognition, 1998.
[16] J. Okamoto, M. Milanova, P. E. Almeida, and M. G. Simoes, “Applications of Cellular Neural Networks for Shape from Shading Problem”, 1999.

[17] S. Gutta and H. Wechsler, "Face recognition using hybrid classifiers", 1997.
[18] A. Pentland, T. Starner, N. Etcoff, A. Masoiu, O. Oliyide, and M. Turk, "Experiments with eigenfaces", In Looking at People Workshop, International Joint Conference on Artificial Intelligence 1993, Chamberry, France, 1993.
[19] R. Chellappa, C.L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey", Proceedings of the IEEE, 83(5): 705–740, 1995.
[20] X. Jia Extending the Feature Set for Automatic Face Recognition. University of Southampton, 1993.
[21] Математический энциклопедический словарь. Под редакцией Ю. В. Прохорова. М.: Сов. Энциклопедия, 1988.847 с. ил.
[22] M. Shah P. S. Tsai. Shape from Shading Using Linear Approximation. University of Central Florida, 1992.
[23] Уоссермен Ф. Нейрокомпьютерная: Теория и практика. http://www.neuropower.de/neurotech.zip, 1992.
[24] Looney C. G.: "Stabilization and Speedup of Convergence in Training Feed-forward Neural Networks", Neuro-computing, 10, 1996, pp. 7-31.
[25] Fahlman S. E.: "An Empirical Study of Learning Speed in Back-Propagation Networks", Technical Report CMU-CS-88-162, Carnegie Mellon University, Pittsburgh PA, June 1988.

تاريخ ورود البحث إلى مجلة جامعة دمشق2004/3/25.