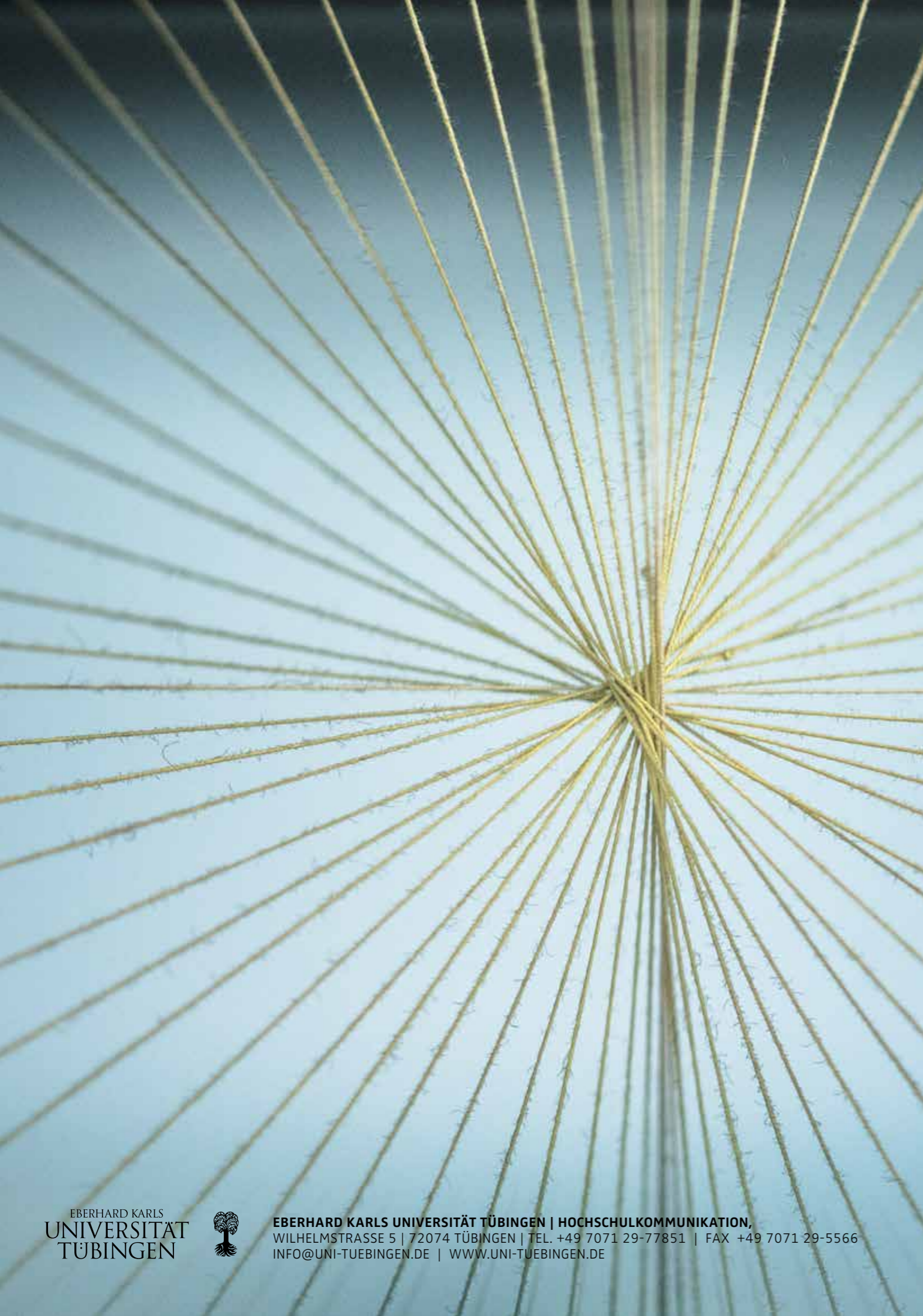


ATTEMPTO!

AUSGABE ISSUE → 45 | 2017
FORUM DER UNIVERSITÄT TÜBINGEN
UNIVERSITY OF TÜBINGEN MAGAZINE



**FORSCHER BRINGEN
MASCHINEN DAS SEHEN BEI
RESEARCHERS TEACH
MACHINES TO SEE**

**MATHEMATIK ERSCHLIESST
NEUE WELTEN
DISCOVERING NEW WORLDS
WITH MATHEMATICS**

**WIE HORMONE UNSER
ESSVERHALTEN BEEINFLUSSEN
HOW HORMONES AFFECT
OUR EATING HABITS**

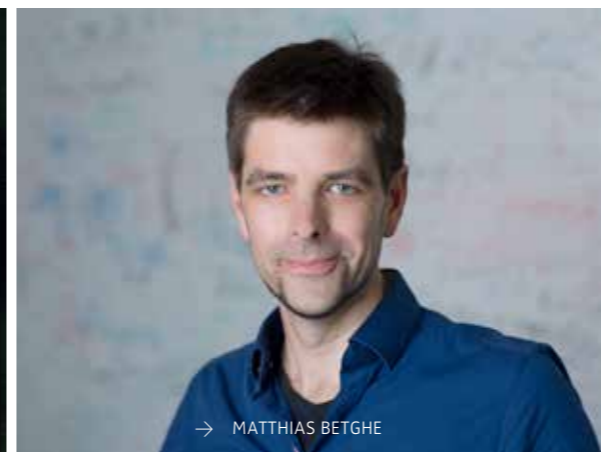


DIE SEHENDE MASCHINE TEACHING MACHINES TO SEE

Künstliche neuronale Netzwerke beherrschen einige Bereiche der Bilderkennung bereits perfekt – und doch ist das Sehvermögen des Menschen noch weit überlegen. Zwei Tübinger Forscher verfolgen aus unterschiedlichen Perspektiven das gleiche Ziel: künstlichen Intelligenzen das robuste Sehen beizubringen.

Artificial neural networks have already mastered some areas of image recognition and yet human vision remains far superior. Two researchers from the University of Tübingen are pursuing the same goal from different perspectives: to teach machines how to see like humans. > *continued on page 13*





anstaltet die Maschinenlern-Community viele Wettbewerbe. So spielte die berühmte ImageNet Challenge eine entscheidende Rolle, um Wissenschaftler von der Leistungsfähigkeit der DNNs zu überzeugen. In diesem Wettbewerb werden Programme trainiert, Objekte auf Fotos zu benennen. Eine Million Fotos „verschlingen“ die Algorithmen während ihres Trainings – bis sie schließlich in der Lage sind, entsprechende Objekte auch in neuen Bildern richtig zu erkennen. Sie erkennen die Katze als Katze, egal in welcher Pose; sie benennen die Anordnung gelb-grüner Bildpunkte als ein Büschel Bananen, auch wenn die Früchte kreuz und quer liegen und halb von Blättern bedeckt sind. In DNNs wird an jeder Stelle des Bildes schrittweise die Information von immer größeren Bildbereichen zusammengefasst, um diese Generalisierungsleistung zu erreichen. Diese sogenannte „retinotop hierarchische Verarbeitung“ findet man auch im visuellen Kortex von Säugetieren. „Echte Gehirne sind natürlich unser Vorbild, unsere Inspiration“, sagt Bethge. DNNs sind bereits heute zu spektakulären Leistungen in der Lage. Ein DNN, das auf dem ImageNet-Datensatz trainiert wurde, lässt sich mithilfe von kleinen Modifikationen auch für ganz andere Aufgaben benutzen. Zum Beispiel hat Matthias Bethge mit Mitarbeitern eine Methode entwickelt, die den Stil eines Kunstbildes erfassen und auf beliebige Fotos übertragen kann. Die Tübinger Neckarfront im Stile von van Goghs Sternennacht etwa? Kein Problem, wie man selber auf der Website <http://deepart.io> ausprobieren kann.

Künstliche Netzwerke lassen sich noch leicht übertölpeln

Und doch funktionieren die Algorithmen im Detail anders als biologische Nervensysteme. Sie haben ihre eigenen verschlungenen Wege gefunden, das Problem des Erkennens anzugehen. So komplex, dass selbst ihre Programmierer oft nicht mehr nachvollziehen können, wie genau sie arbeiten. Felix Wichmann nennt das augenzwinkernd „Cowboyland“ – eilig steckten KI-Forscher ihre Claims auch in unbekanntem Terrain ab. Für ihn basiert der große mediale Hype um das sogenannte Deep Learning der künstlichen neuronalen Netzwerke denn auch auf einer falschen Annahme: „Oft wird gesagt, die Netzwerke funktionierten ganz ähnlich wie das menschliche Sehsystem, es gäbe klare Parallelen.“ Der Psychophysiker ist skeptisch.

Er führt die Tricks an, mit denen die Kunsthirne noch einfach verwirrt werden können. Auf sogenannten „Salt and Pepper Noise“ etwa, also auf eingestreute schwarze und weiße Stör- →

Durch das menschliche Gehirn inspiriert

„Ich will zunächst die Informationsverarbeitung von intelligenten Systemen verstehen, wie viele andere Neurowissenschaftler auch“, erklärt Matthias Bethge. Er verortet sich inhaltlich zwischen Künstlicher Intelligenz (KI) und Neurowissenschaft, im Bereich des maschinellen Lernens. „Die Herausforderung im Maschinenlernen ist aber, nicht nur zu verstehen, wie ein Mechanismus funktioniert, sondern tatsächlich auch neue Systeme zu bauen.“ Seine Arbeitsgruppe erstellt hoch komplexe, weit verschachtelte Computeralgorithmen: sogenannte Deep Neural Networks (DNN), also tiefe, künstliche neuronale Netze aus vielen hintereinandergeschalteten Signalverarbeitungselementen. Ihre grundsätzliche Architektur ist vorgegeben, die genaue Funktion der einzelnen Elemente jedoch dynamisch – ähnlich wie beim biologischen Nervennetz werden die Signalflüsse immer wieder angepasst, um die eingehenden Daten immer besser zu verarbeiten. Die Gewichtungen von Millionen Verknüpfungen variieren und reagieren so lange auf korrigierendes Feedback, bis das neuronale Netz eine Aufgabe optimal lösen kann.

Um die Leistungsfähigkeit von verschiedenen Lernalgorithmen vergleichen zu können, ver-

- 01 NOCH IST DAS MENSCHLICHE SEHVERMÖGEN DEM KÜNSTLICHER NEURONALER NETZWERKE WEIT ÜBERLEGEN.
- 02 DIE WISSENSCHAFTLER HABEN EINEN ALGORITHMUS ENTWICKELT, MIT DEM SICH DER STIL EINES KUNSTWERKS AUF JEDES BELIEBIGE FOTO ÜBERTRAGEN LÄSST.
- 01 HUMAN VISION REMAINS FAR SUPERIOR TO ARTIFICIAL NEURAL NETWORKS.
- 02 THE SCIENTISTS DEVELOPED AN ALGORITHM, WHICH TRANSFERS STYLISTIC ELEMENTS OF AN ARTWORK TO A PHOTOGRAPH.

> deutsch

// _____ Auf unser Sehvermögen ist Verlass. Augenblick für Augenblick treffen Millionen von Photonen auf die Netzhaut der Augen. Unser Gehirn schafft es mühelos, Irrelevantes aus dieser Datenflut zu filtern und eine Bedeutung aus den Bildinformationen zu konstruieren. Ob bei schlechtem Licht oder leichtem Nebel, aus der Vogelperspektive oder halb verdeckt, wir sehen und verstehen: Apfel oder Gesicht, Freund oder Feind. Der Prozess ist äußerst störungsarm, robust – und hoch komplex. Bis heute ist er nur ansatzweise verstanden. Das zeigt sich besonders dann, wenn Wissenschaftler versuchen, unsere visuelle Wahrnehmung im Modell nachzubauen. Wenn Maschinen lernen sollen, Bilder zu verstehen. In manchen Fällen klappt das hervorragend. Die Leistungsfähigkeit von künstlichen neuronalen Netzen und mit ihnen das Maschinenlernen hat sich in den letzten fünf Jahren dramatisch verbessert; bei bestimmten Aufgaben können moderne Algorithmen Bilder schon besser kategorisieren als menschliche Versuchspersonen. Doch dann wiederum lassen sich die Maschinen durch einfachste Tricks übertölpeln.

Matthias Bethge und Felix Wichmann haben sich beide zur Aufgabe gemacht, den Sehprozess genauer zu durchschauen. Doch gehen sie das Problem aus unterschiedlichen Richtungen an: Der Physiker und Neurowissenschaftler Bethge leitet die Ar-

beitsgruppe Computational Neuroscience and Machine Learning am Exzellenzcluster „Centre for Integrative Neuroscience (CIN)“ der Universität Tübingen. Als Experte für maschinelles Lernen nimmt er den Computer als Ausgangspunkt; er nutzt künstliche neuronale Netze, um das Sehen zu ergründen. In unzähligen Durchläufen wurden sie trainiert, um ganz bestimmte Aufgaben zu erlernen. Wichmann dagegen geht vom Menschen aus. Er leitet den Tübinger Lehrstuhl für Neuronale Informationsverarbeitung; als Psychophysiker beschäftigt er sich mit psychologischen Prozessen, die der menschlichen visuellen Wahrnehmung zugrunde liegen, zerlegt sie in kleinste Schritte und protokolliert in seinen Modellen die genaue Architektur unseres Sehvermögens. Beide sind sie Teil des Sonderforschungsbereichs „Robustheit des Sehens“ (SFB 1233), der seit Anfang 2017 an der Uni Tübingen besteht. Beide sind sie Bewohner des neu gegründeten „Cyber Valley“ (siehe Kasten). Und beide treffen sich, so das Ideal, irgendwann in der Mitte – zwischen Mensch und Maschine.

pixel im Bild, reagieren Bilderkennungsalgorithmen empfindlich, oft nennen sie dann eine falsche Kategorie. „Die Fehler sind so teilweise unmenschlich“, sagt Wichmann. „Auch mit einem halben Kilo Salz und Pfeffer auf einem Katzenfoto wird ein Mensch niemals sagen: ‚aha, ein Haus.‘“ Ein weiterer elementarer Unterschied zur Maschine liegt für ihn in der enormen Anpassungsfähigkeit der menschlichen Wahrnehmung. Ohne Training sind wir in der Lage, Dinge zuverlässig in völlig neue Kategorien einzusortieren. Felix Wichmann lässt den Blick durch sein Büro schweifen. „Sie könnten mir zum Beispiel alle Gegenstände nennen, die als Gefäß dienen könnten. Oder all jene, die sich beim Anfassen kalt anfühlen.“ Diese spontane Transferleistung würde einen Algorithmus überfordern.

Zwei unterschiedliche Ansichten gibt es laut Wichmann über die künstlichen Sinnesnetzwerke. Maschinenaffine Optimisten glauben, man müsse nur noch kleine Verbesserungen vornehmen. Dann sei das menschliche Vorbild erreicht. Die Skeptiker sehen noch eine riesige Lücke. „Meine Meinung ist: Wir haben mit den DNNs tolle technische Tools“, sagt der Wissenschaftler. „Aber noch lange nicht die Einsicht, was der Mensch genau macht im Prozess der Wahrnehmung.“

Kombination

zweier Forschungsansätze

Um zu dieser Einsicht zu gelangen, schickt Felix Wichmann Menschen in die Höhle. Komplett dunkel ist es in seinem Psychophysiklabor: schwarze Farbe an Wänden, Decke und Boden; schwarze Samtvorhänge, schwarze Stühle. Kein Lichtstrahl soll die aufwendigen Experimente stören, mit denen Probanden hier stunden-, tage-, manche sogar jahrelang beschäftigt sind. Wie reagieren die Probanden auf geringste Kontrastunterschiede? Welche Abstufungen kann das menschliche Sehvermögen erkennen, wie verändern diese den ganzen Wahrnehmungsprozess? Wichmann fängt unten an, ganz im Detail, bei der sogenannten „early vision“, den frühen Stufen des Sehprozesses. Und nutzt seine Erkenntnisse, um sie in Computermodellen nachzubauen. Die sind zwar noch längst nicht so komplex wie die künstlichen neuronalen Netze seines Kollegen – bisher bestehen sie aus zwei Ebenen –, „dafür haben wir sie bis aufs kleinste Detail verstanden“, versichert Wichmann.

Eine Idee ist nun, die beiden Ansätze zu kombinieren. Die ersten beiden Ebenen eines Seh-Netzwerks etwa könnten aus Wichmanns biologisch nachempfundenen Algorithmen bestehen, ergänzt um die Bethge'schen DNNs. Die menschlichere Maschine sozusagen – ein bestechender Gedanke.

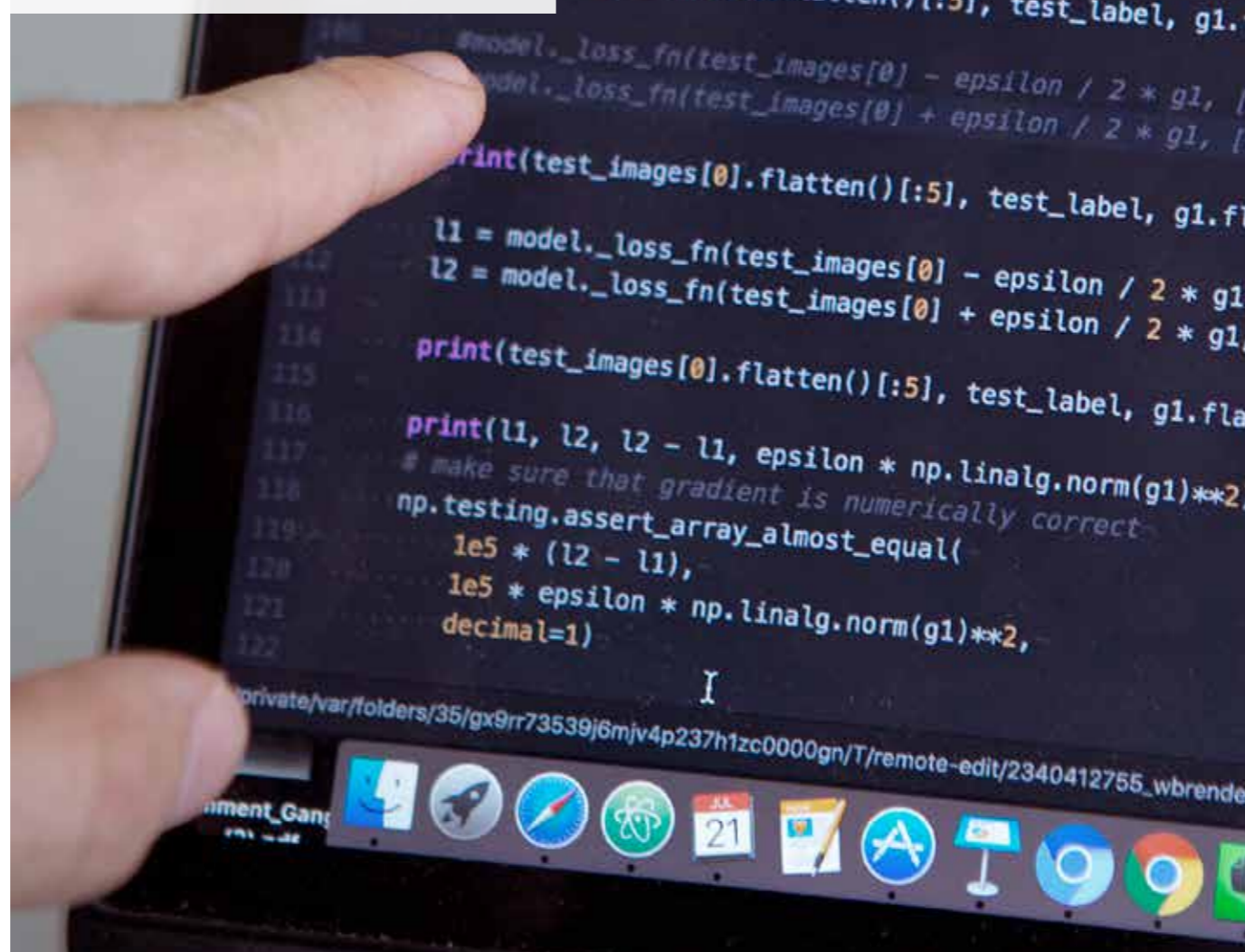
Es herrscht nicht immer nur traute Forscherharmonie zwischen den beiden Wissenschaftlern – obwohl sie privat befreundet sind und sich gut leiden können. Im Gegenteil, es ist gerade die entstehende Reibung, die sie aneinander zu schätzen wissen. Der eine mosert über unüberschaubar gewordene künstliche Netzwerke, der andere über rudimentär anmutende Versuche mit Streifenmustern, erzählen sie. So haben sie sich gut eingependelt – in einer durchaus kritischen, aber produktiven Kooperation. Einig sind sie, dass die Frage nach der Robustheit des Sehens die entscheidende ist, an der künstliche Systeme noch viel lernen können. Bethges Arbeitsgruppe hat einen Wettbewerb entwickelt (<http://robust.vision>), der im Feld des maschinellen Lernens ermuntern soll, robustere Algorithmen zu entwickeln. Für diesen wurden in Tests in Wichmanns Labor gemessen, wie viel robuster sich menschliches Sehen gegenüber dem maschinellen derzeit darstellt. Spätestens wenn es darum geht, in Zukunft etwa autonom fahrende Autos durch eine komplexe Welt zu steuern, ist diese Frage wichtig – durch die gemeinsame Forschung kann es den Tübinger Kollegen vielleicht gelingen, maschinelles Sehen hier an das menschliche anzunähern. ____//

DAS CYBER VALLEY: DIE KÜNSTLICHE INTELLIGENZ FINDET EIN ZENTRUM

> Seit 2016 liegt Tübingen im Cyber Valley: einem neuen Hotspot für die Erforschung künstlicher Intelligenz. Partner aus Wissenschaft und Industrie bauen in dem Projekt gemeinsam ihre Expertise auf den Gebieten maschinelles Lernen, Robotik und Computer Vision aus. Kerngedanke ist neben dem Aufbau eines weltweit renommierten Forschungszentrums der Transfer in die Wirtschaft – das Cyber Valley soll ein innovatives Klima schaffen, das kreative Start-ups fördert und mit etablierten Unternehmen vernetzt. (www.cyber-valley.de)

> An der „International Max Planck Research School (IMPRS) for Intelligent Systems“ werden Doktoranden ausgebildet, an den Universitäten Tübingen und Stuttgart insgesamt zehn neue Professuren eingerichtet. Auch im neuen Sonderforschungsbereich zur Robustheit des Sehens arbeiten Forscher des MPI für Intelligente Systeme und der Universität Tübingen gemeinsam (<http://robust.vision/CRC1233>).

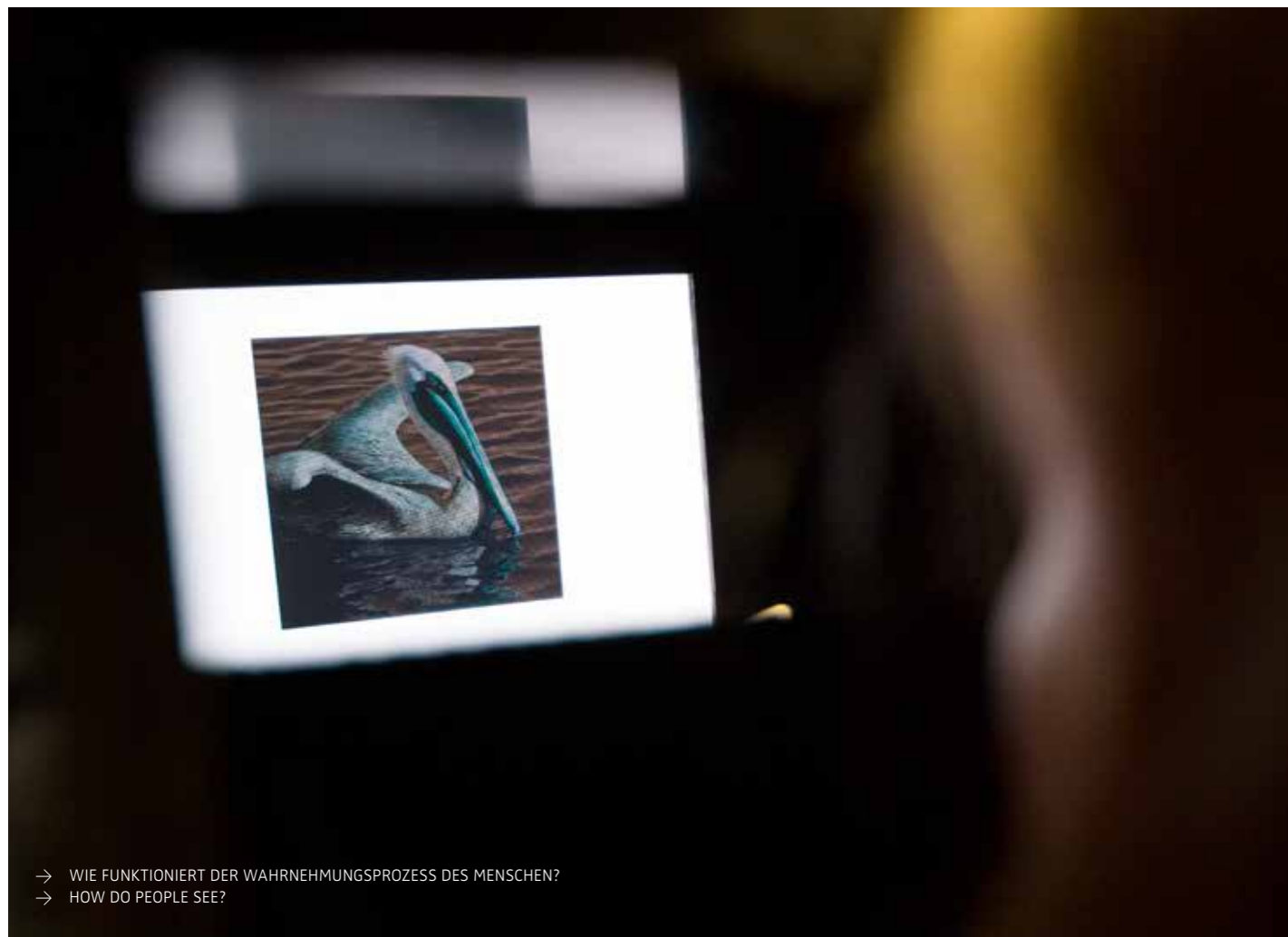
> Neben dem Land Baden-Württemberg als Hauptförderer sind die Universitäten Tübingen und Stuttgart sowie das Max-Planck-Institut für Intelligente Systeme mit seinen Standorten Tübingen und Stuttgart am Cyber Valley beteiligt; als Industriepartner fungieren BMW, Facebook, Daimler, Porsche, Robert Bosch und ZF Friedrichshafen. In den kommenden fünf Jahren finanzieren alle Beteiligten das Projekt mit einem hohen zweistelligen Millionenbetrag.



> englisch

//_____ Every second, millions of photons hit the retina and are transformed into meaning by our brains. This process is remarkably robust – and highly complex even though it usually appears to be effortless. Even under challenging conditions such as low light or fog, unfamiliar perspectives or occlusion, our brains are extremely effective to make sense of the visual input. In specific cases, machines are already able to compete with humans. The performance of artificial neural networks and machine learning dramatically improved over the past five years, and in certain tasks modern algorithms can even surpass human performance. Nevertheless, machines can still be easily fooled with simple tricks.

Matthias Bethge and Felix Wichmann have both set out to take a close look at the human visual process, but from two different perspectives. Physicist and neuroscientist Bethge heads the Computational Neuroscience group in the Centre for Integrative Neuroscience (CIN) at the University of Tübingen. As an expert in machine learning, he is using artificial neural networks to investigate visual information processing. These networks are trained with large amounts of data to learn specific vision tasks. Wichmann's approach in turn is more focused on obtaining accurate descriptions of human vision. As a psychophysicist and professor of Neural Information Processing, he deals with the psychological processes that underlie human visual perception, dividing them into individual steps, and reconstructing the exact architecture of our vision in his models. Both scientists are members of the collaborative research center "Robust Vision" (SFB 1233), which was founded at the University of Tübingen in early 2017, and they are also both based in the new Cyber Valley (see boxout). →



→ WIE FUNKTIONIERT DER WAHRNEHMUNGSPROZESS DES MENSCHEN?
→ HOW DO PEOPLE SEE?

Artificial networks can still
be fooled easily

Algorithms in machine neural networks work differently to biological neural networks. They have found their own intricate ways of addressing the problem of recognition. Complex enough that even their programmers do not know exactly how they work. Felix Wichmann describes this as “Cowboy land” and notes that many AI researchers have been quick to make somewhat overstated claims in unknown terrain. For Wichmann, the media hype surrounding deep learning in artificial neural networks is at least partially based on a false assumption: “It is often said that artificial networks function like the human visual system and that there are clear algorithmic parallels.” However, as a psychophysicist Wichmann is skeptical and lists several tricks which can easily confuse artificial intelligence. For example, algorithms are still much more sensitive to “salt and pepper noise” or a black and white interference pattern in an image which causes them to categorize images incorrectly. Another fundamental difference to machines lies in the extensive adaptability of human perception. Without training, we can sort things reliably and on the fly into completely new categories.

According to Wichmann, there are two different views on artificial neural networks. Machine learning optimists believe that only small improvements must be made to reach levels of human performance. Skeptics believe that there is still a huge gap. “DNNs are great technical tools”, says Wichmann, “but they are far from demonstrating how human visual processing works.”

Combining research
approaches

Participants in Felix Wichmann’s psychophysics lab enter a world of darkness with black walls, ceilings, floors, curtains, and chairs. The intention is to block extraneous light from interfering with the visual experiments which participants undertake over hours, days, and sometimes even years. How do the subjects respond to the slightest differences in contrast? What are the minimal differences that can be recognized by humans and how do they affect perception? Wichmann starts bottom-up with “early vision” by reconstructing early stages of human visual processing from his findings using computer models. Although they are far less complex than the artificial neural networks used by Bethge – so far, they consist of two levels – Wichmann assures that his team can understand each element to the most minute detail.

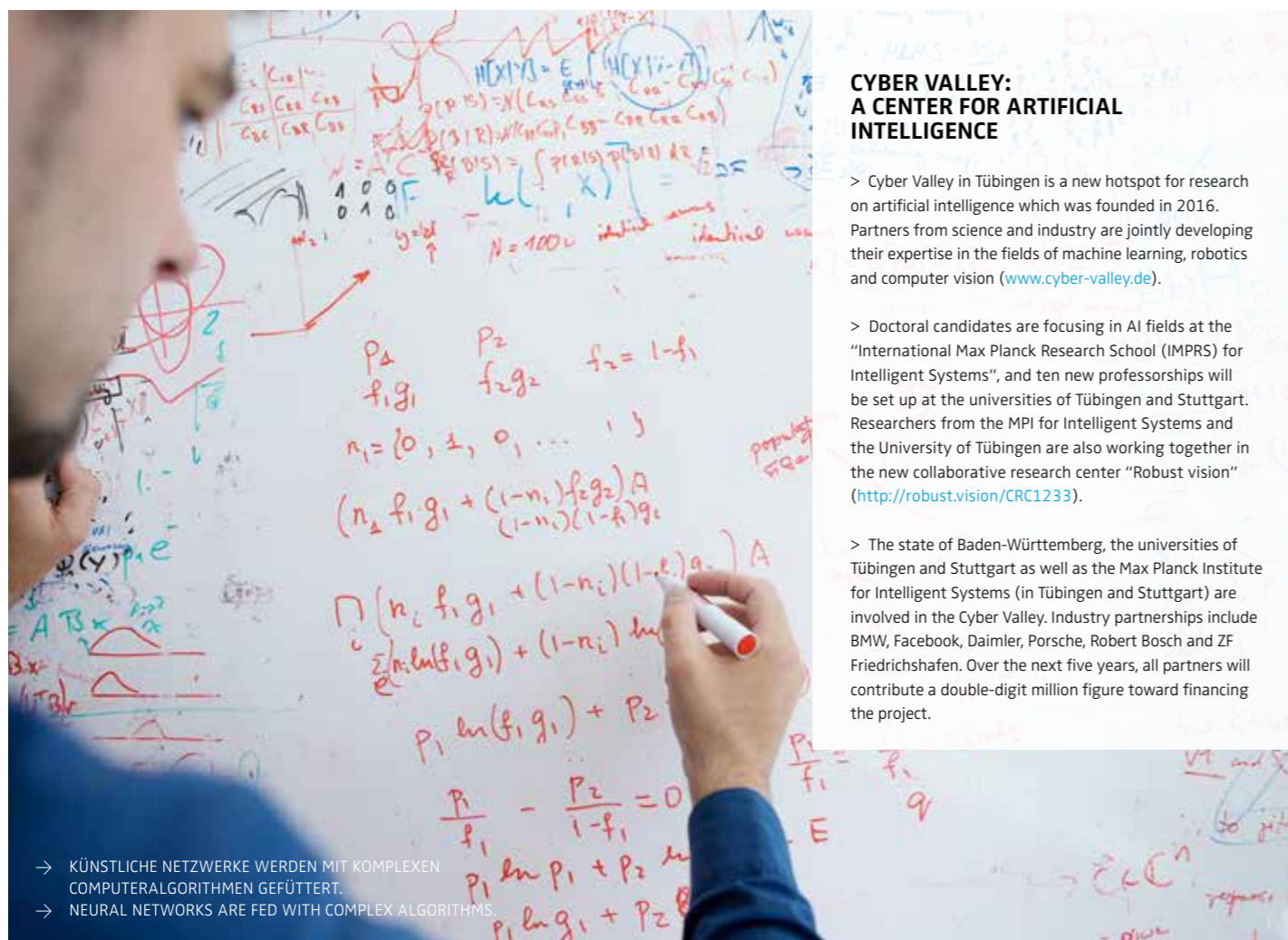
One future goal is to combine the two approaches to make machine vision more robust. Bethge’s group launched a competition (<http://robust.vision/benchmark>) to encourage the field to make neural networks more resilient against image perturbations such as the “salt and pepper noise” that is studied and compared against human vision in Wichmann’s lab. The combination of Wichmann’s biologically simulated algorithms and Bethge’s artificial deep neural networks could be one prospective solution to this challenge. More generally, a precise understanding of the robustness of machine recognition will definitely become highly important in the future, if cars should drive autonomously through complex environments. The collaboration between these colleagues in Tübingen could make an important contribution to tackle this question and make machine vision as robust as human vision. ____//

Inspired by
the human brain

“I want to primarily understand how intelligent systems process information, like many other neuroscientists”, explains Matthias Bethge. He describes his research as positioned between artificial intelligence (AI) and neuroscience in the field of machine learning. “However, the challenge in machine learning is not just to understand how a mechanism works, but actually to build new systems.” Bethge’s group creates highly complex computer algorithms: deep neural networks (DNN), which are complex compositions of many successive signal processing elements. Their basic architecture is fixed but the exact function of the individual elements is dynamically adjusted – like biological neural networks, the signal flow adapts to process incoming data with greater efficiency. The weights of millions of connections vary and react to corrective feedback until the neural network can optimally solve a task.

To be able to compare the performance of different algorithms, the machine learning community holds many competitions. A famous one is the ImageNet Challenge, which played a decisive role in convincing scientists of the power of DNNs. In this competition, programs are trained to name objects in photos. The algorithms process one million images during training until they are finally able to correctly recognize objects in new images. In DNNs, information from increasingly larger image areas is grouped step by step at every point in the image to achieve generalizations. This feature of “retinotopic hierarchical processing” is also found in the visual cortex of mammals. “Real brains serve as our model and inspiration”, says Bethge.

DNNs are already capable of spectacular performance. A DNN that has been trained on the ImageNet dataset can also be used for quite different tasks with the help of small modifications. For example, Matthias Bethge and his team have developed a method which can transfer the stylistic elements from an artwork to a photograph. Have you ever wondered what the Neckar waterfront in Tübingen would look like in the style of Vincent van Gogh’s ‘Starry Night Over the Rhône’? Try it out at <http://deepart.io>.



→ KÜNSTLICHE NETZWERKE WERDEN MIT KOMPLEXEN
COMPUTERALGORITHMEN GEFÜTTERT.
→ NEURAL NETWORKS ARE FED WITH COMPLEX ALGORITHMS.

**CYBER VALLEY:
A CENTER FOR ARTIFICIAL
INTELLIGENCE**

> Cyber Valley in Tübingen is a new hotspot for research on artificial intelligence which was founded in 2016. Partners from science and industry are jointly developing their expertise in the fields of machine learning, robotics and computer vision (www.cyber-valley.de).

> Doctoral candidates are focusing in AI fields at the “International Max Planck Research School (IMPRS) for Intelligent Systems”, and ten new professorships will be set up at the universities of Tübingen and Stuttgart. Researchers from the MPI for Intelligent Systems and the University of Tübingen are also working together in the new collaborative research center “Robust vision” (<http://robust.vision/CRC1233>).

> The state of Baden-Württemberg, the universities of Tübingen and Stuttgart as well as the Max Planck Institute for Intelligent Systems (in Tübingen and Stuttgart) are involved in the Cyber Valley. Industry partnerships include BMW, Facebook, Daimler, Porsche, Robert Bosch and ZF Friedrichshafen. Over the next five years, all partners will contribute a double-digit million figure toward financing the project.