

Real Time Image Upscaling & Deep Learning

Michael Singer

Hochschule der Medien, Stuttgart, Deutschland

Wintersemester 22/23 · AM3 · Aktuelle Themen bei Professor Dr. Koch

10.03.2023



1. Abstract

Deep Learning Super Sampling und andere Formen des *Realtime-Image Upscalings* bieten das Potential der Einsparung von Rechenleistung durch KI-gestütztes Image Upscaling, während verhältnismäßig geringe Einbußen in Form der wahrgenommenen Bildqualität zu verzeichnen sind. Das Verfahren wird durch speziell neuronale Netze ermöglicht, welche Bilder mit geringerer Auflösung auf eine höhere Auflösung hochskaliert. Die Schnittstelle hierfür bieten unter anderem *Machine Learning*¹ Rechenkerne, welche das Upscaling in Echtzeit ermöglichen.

2. Einleitung

Bei *Image Upscaling* handelt es sich um den Prozess des Hochskalierens von Medieninhalten mit einer geringeren Auflösung auf eine höhere Auflösung (Salian, 2020). Es gibt eine Vielzahl an Methoden des Image Upscalings, welche im weiteren Verlauf dargestellt werden, wobei vor allem der Fokus auf Upscaling m.H. von *Deep Learning*² Technologien liegt. Dieser Prozess wird in verschiedensten Branchen eingesetzt, darunter die Medizin, sowie die Video-spielindustrie. Im Folgenden werden die Anwendungen, Methoden und das Potential, sowie Herausforderungen des Image Upscalings mit Hilfe von Deep Learning dargestellt.

¹ *Machine Learning* beschreibt bei KI das Trainieren des künstlichen neuronalen Netzes und die dadurch von diesem Netz erworbene Fähigkeit, der vorliegenden Daten Regelmäßigkeiten oder anderweitig neue Informationen zu erkennen (Murphy, 2012).

² *Deep Learning* beschreibt eine Klasse von Machine Learning Algorithmen, die mehrere Schichten verwendet, um Merkmale auf höherer Ebene aus dem rohen Input zu extrahieren (Deng, 2014).

3. Notwendigkeit von Image-Upscaling

Zunächst gilt festzuhalten, weshalb die Technologie des Image Upscalings Relevanz in der heutigen Zeit hat. Die Notwendigkeit stammt daher, dass Displays und anderweitige Wiedergabegeräte, wie beispielsweise VR-Brillen den Zugang zu immer höheren Auflösungen zu verhältnismäßig erschwinglichen Preisen ermöglichen. Demnach haben im Jahr 2021 44% aller Haushalte in den USA einen 4K-Fernseher, verglichen zu 31% im Jahr 2019 (Laricchia, 2022). Die stetig wachsende Verdrängung von Full-HD durch 4K oder höheren Auflösungen führt dementsprechend zu einer wachsenden Berechnungslast, sowie der Voraussetzungen, dass Inhalte in der nativen Auflösung bereitgestellt werden können. Gerade im Bereich TV, Streaming und Live-Übertragungen ist das jedoch nicht immer der Fall.

An dieser Stelle werden Upscaling-Algorithmen eingesetzt, um beispielsweise einen Full-HD Livestream auf einem 4K-Fernseher wiedergeben zu können. Wie in Abbildung 1 zu sehen, gibt es einen deutlichen Sprung in der Anzahl der Pixel bei fortschreitender Auflösung, was sich in der benötigten Übertragungsbandbreite für Streaming-Medien, sowie der Rechenlast bei interaktiven Medien widerspiegelt.

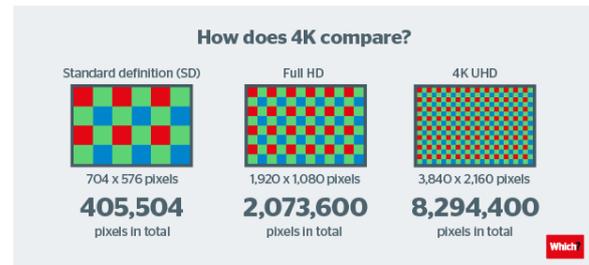


Abbildung 1: Vergleich der Pixelanzahl von verschiedenen gängigen Bildschirmauflösungen (Pratt, 2023)

Demnach sei die Notwendigkeit von Image Upscaling durch ein Ungleichgewicht der möglichen Auflösung und der tatsächlich umsetzbaren Auflösung gegeben. Die Limitationen sind hierbei unter anderem: Mögliche Übertragungsbandbreite, verfügbare Rechenleistung, sowie die Auflösung von Medieninhalten.

4. Methoden und Entstehung von Image Upscaling

Um Image Upscaling mit Hilfe von Deep Learning Algorithmen zu verstehen, müssen zunächst grundlegende Image-Upscaling Methoden aufgeführt werden. Hierbei kann festgehalten werden, dass das Ziel darin besteht, ein Bild mit geringerer Auflösung auf ein größeres Display, bzw. eine höhere Bildschirmauflösung zu übertragen. Hierbei werden die Pixel des Bildes mit der niedrigeren Auflösung kopiert und wiederholt, damit alle Pixel des Displays mit höherer Auflösung gefüllt werden können.

Anschließend wird das Bild mit Hilfe von Anti-Aliasing-Filtern geglättet, um dem dadurch entstehenden „Treppcheneffekt“ (Aliasing) entgegenzuwirken (siehe Abb. 2).



Abbildung 2: Darstellung von Aliasing und Anti-Aliasing (Chaitanya, 2021).

4.1 Anti-Aliasing

Beim Anti-Aliasing gibt es einige verschiedene Methoden der Kantenglättung. Dieser Effekt tritt besonders bei geringeren Auflösungen auf – demnach führt der Schritt in Richtung höherer Pixeldichte zu weniger kantigeren Bildern – ist jedoch verbunden mit höherer Rechenlast. Besonders im Bereich des Gamings gibt es einige verschiedene Anti-Aliasing Verfahren, welche sich in der zu erreichbaren Bildqualität, sowie der benötigten Rechenleistung unterscheiden. Zum einen gibt es das sog. Multisample Anti-aliasing (MSAA). Hierbei handelt es sich um die einfachste Form des Anti-Aliasing, wobei hier eine Kombination aus Kantenerkennungsalgorithmen und Pixeladdition vorgenommen wird. Bei Temporal Anti-Aliasing (TXAA) handelt es sich um einen fortschrittlicheren und

verhältnismäßig neueren Anti-Aliasing Algorithmus, welcher MSAA und spezielle Filter innerhalb und außerhalb von Pixeln verwendet, um ein homogeneres Bild als nur mit Hilfe von MSAA zu erzeugen. Dies geschieht jedoch auf Kosten der Rechenleistung (Stanar, 2020). Fast Approximate Anti-Aliasing (FXAA) wurde von NVIDIA entwickelt und stellt einer der beliebtesten Anti-Aliasing Methoden dar. Hierbei ist vor allem der minimale zusätzliche Rechenaufwand hervorzuheben, den dieses Verfahren verursacht. Es fungiert als eine Art pixelbasierter Weichzeichner, der die kantigen Linien einfängt und das Bild insgesamt etwas weicher wirken lässt.

Super Sampling Anti-Aliasing (SSAA) stellt einen der ersten Anti-Aliasing Filter dar. Hierbei wird als Basis der Kantenglättung ein höher aufgelöstes Bild verwendet, welches auf die gewünschte Zielauflösung heruntergerechnet wird. Dies führt denkbarerweise zu einer deutlich höheren Berechnungslast.

Das Pendant hierzu, also Super Sampling auf Basis eines Bildes, das eine niedrigere Auflösung besitzt als die Zielauflösung stellt den Fokus dieses White Papers dar. NVIDIA präsentierte hierfür im Jahre 2019 ihren Deep Learning Super Sampling (DLSS) Upscaling Algorithmus.

4.2 Image-Upscaling mit Hilfe von Deep Learning Technologien

Das Upscaling mit Hilfe von AI, bzw. eines Deep Learning Netzwerks unterscheidet sich deutlich von der Herangehensweise des traditionellen Upscalings. Bei einem Bild mit niedriger Auflösung sagt ein Deep-Learning-Modell ein hochauflösendes Bild voraus, das so verkleinert wird, dass es wie das ursprüngliche, niedrig aufgelöste Bild aussieht (Salian, 2020). Um einen möglichst sauberen Übergang zwischen verschiedenen Frames zu ermöglichen, wird bei dieser Upscaling Methode auch auf Parameter von vorherigen Frames zugegriffen, sog. „*Temporal Feedback*“. Eine Schlüsselkomponente hierfür sind sogenannte *Motion-Vektoren*, welche eine Repräsentation der Bewegung von Objekten in einer Szene über die Zeit sind (Davenport et al., o. J.). Sie werden sowohl zur Videokompression, als auch bei Image Upscaling Verfahren (mit Hilfe von neuronalen Netzen) verwendet, um qualitativ hochwertige, hochskalierte Bilder aus den Eingabebildern mit geringerer Auflösung zu rekonstruieren. Bei der Verwendung von Motion-Vektoren geht es auch darum, Artefakte, wie beispielsweise „Ghosting“ vorzubeugen. Hierbei handelt es sich um „Nachziehen“, bzw. „Verschlierung“ von Bildinhalten aufgrund von Pixeln, die nicht rechtzeitig aktualisiert werden, wodurch Teile des Bildes verschwimmen können.

4.3 Traditionelle Image-Upscaling - Methoden

Im Gegensatz zu Upscaling-Methoden, welche auf neuronalen Netzwerken beruhen gibt es auch traditionelle Möglichkeiten des Upscalings. Diese benötigen keinen Zugriff auf Trainingsdaten, oder Informationen der vorherigen Frames, wie beispielsweise Jitter-Offsets oder Motion-Vektoren (Syed, 2021).

Zum einen gibt es die Möglichkeit der „*Bilinear Interpolation*“. Hierbei handelt es sich um die simpelste Methode des Upscalings, bei der Werte von fehlenden Pixeln durch eine lineare Interpolation zwischen benachbarten Pixeln berechnet werden. Diese Methode ist verhältnismäßig schnell und einfach zu implementieren, jedoch kann es zu einem Verlust von Details und Schärfe im Bild führen.

Bei „*Bicubic Interpolation*“ handelt es sich um eine Verbesserung der „*Bilinear Interpolation*“ bei der der Wert von fehlenden Pixeln durch eine kubische Interpolation zwischen benachbarten Pixeln berechnet wird. Diese Methode kann zu einem schärferen und detaillierteren Bild führen als die „*Bilinear Interpolation*“, jedoch kann es dabei zu einem höheren Rechenaufwand kommen.

4.4 Vergleich der Image Upscaling-Methoden

In Abbildung 3 ist anhand der Illustration einer seismologischen Darstellung des Golfstroms ein deutlicher Unterschied in Bezug auf die Schärfe festzustellen.

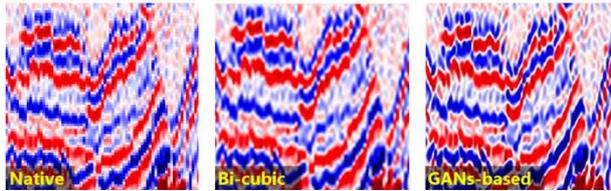


Abbildung 3: Vergleich zwischen bicubic und GAN³-based Image Upscaling (Lu, 2018).

GANs können mit extrem vielen hochauflösenden Bildern speziell auf das Hochskalieren von niedrig aufgelösten Bildern trainiert werden. Hierbei werden Muster der hochauflösten trainierten Bilder erkannt und auf das niedrig aufgelöste Bild angewandt, sodass dieses dem hochauflösten Bild möglichst ähnelt.

Durch diesen Ansatz lässt sich deutlich Rechenleistung einsparen, da effektiv nur ein niedrig aufgelöstes Bild berechnet werden muss (siehe Abb. 5).

Das Hochskalieren dieses Bildes übernimmt das Deep Learning Netzwerk, welches zuvor mit etlichen hochauflösten Bildern trainiert wurde.

In der Praxis bedeutet die Einsparung der Rechenlast, dass diese Technologie besonders im Bereich der interaktiven Medien eingesetzt werden kann.

In Abbildung 5 ist zudem zu erkennen, was der markanteste Unterschied von traditionellen Upscaling-Verfahren und Verfahren mit Hilfe eines Deep-Learning Netzwerks ist. Letztere erzielen bei einem (teils deutlich) geringerem Rechenaufwand ein Bild, welches dem nativ berechnet sehr nahekommt – und unter Umständen sogar schärfer wirkt, als das native Ausgangsbild.

Der Anwendungsbereich von „DLSS“ und anderen Image Upscaling Verfahren mit Hilfe eines Deep Learning Netzwerks liegt unter anderem bei Games, sowie in der Medizin.

5. Super Resolution in Games

Einsparung von Rechenleistung bei nahezu gleichbleibender Bildqualität ermöglicht bei Games vor allem den Einsatz von beeindruckenden grafischen Effekten, bzw. Berechnungsmethoden. Ein Beispiel hierfür ist die Einführung von realtime „Ray Tracing“, welches im Jahr 2018 von NVIDIA vorgestellt wurde. Unter Ray Tracing versteht man eine Rendering-Technik, mit der die Beleuchtung einer Szene und ihrer Objekte realistisch simuliert werden kann, indem physikalisch genaue Reflexionen, Brechungen, Schatten und indirekte Beleuchtung wiedergegeben werden.

³ GAN: Generative adversarial network. Sie stellen einen Ansatz des „generative Modellings“ (selbständige Erkennung von Mustern eines Netzwerks) mit Hilfe von Deep Learning Methoden dar (Brownlee, 2019).

Beim „Ray Tracing“ werden Computergrafikbilder erzeugt, indem der Weg des Lichts von der Sichtkamera (die den Blick auf die Szene bestimmt) durch die 2D-Sichtebene (Pixelebene), in die 3D-Szene und zurück zu den Lichtquellen verfolgt wird (Caulfield, 2018). „Ray Tracing“ an sich ist keine neue Technologie und wurde bereits beim Rendering von 3D-Szenen eingesetzt – die Umsetzung dieser Berechnungsmethode in Echtzeit ist jedoch eine Leistung, zu der „DLSS“ einen wichtigen Beitrag geleistet hat. Der zusätzliche Rechenaufwand von „Ray Tracing“ wird durch den Performancegewinn von „DLSS“ eingefangen, wodurch die Kombination von „Ray Tracing“ und „DLSS“ zu einer höheren Bildqualität bei ähnlicher Bildwiederholrate, wie bei nativer Berechnung ohne „Ray Tracing“ führen kann (siehe Abb. 4).

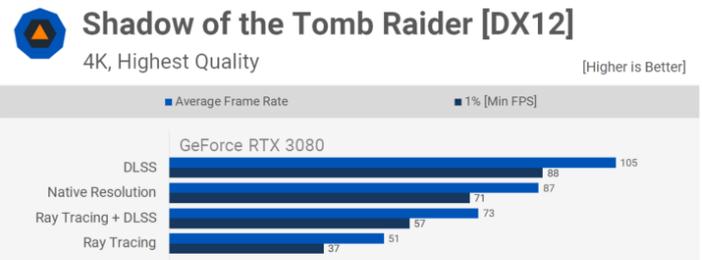


Abbildung 4: Performancevergleich von DLSS, Raytracing und nativer Auflösung (Schuesser, 2020)

Die Einführung von „DLSS“ im Jahre 2018 gab den Startschuss für KI-gestützten Image-Upscaling. In der Zwischenzeit entwickelten sich einige Alternativen von Intel und AMD. AMD stellte mit „Fidelity Super Resolution“ (FSR) im Jahr 2021 ein Äquivalent zu Nvidias „DLSS“ vor. Dieses besteht besonders darin, dass es keine besonderen Hardwareanforderungen besitzt – im Gegensatz zu DLSS. Intel stellte mit „XeSS“ ebenfalls ein Konkurrenzprodukt vor, welches seit 2022 erhältlich ist.

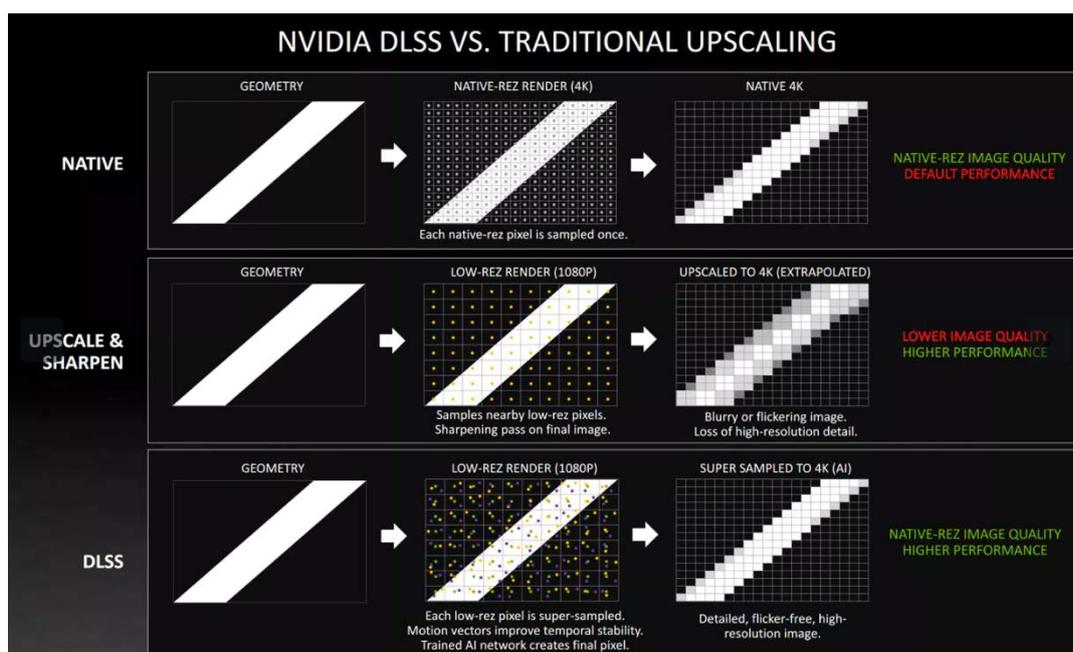


Abbildung 5: Traditionelles Upscaling vs. Nvidia DLSS (Syed, 2021)

5.1 Entwicklung von Image-Upscaling im Gaming

Als Nvidias „DLSS“ im Jahr 2018 zum ersten Mal in Version 1.0 dem Massenmarkt Gaming vorgestellt wurde, gab es noch einige Startschwierigkeiten. So konnte die Image-Upscaling Technologie zum einen auf Performance-Ebene bereits erste Erfolge erzielen, jedoch war die erzielbare Bildqualität der nativen Auflösung deutlich unterlegen. In dieser Version war „DLSS“ noch von starken Artefakten geprägt. In Version 2 konnte NVIDIA deutliche Fortschritte der Bildqualität erzielen und überschritt in dieser Version bereits die Wahrnehmungsgrenze zu nativen Auflösungen bei deutlich besserer Performance. „DLSS 3.0“ ermöglicht das Generieren eigenständiger Frames durch das neuronale Netzwerk, während in den vorherigen Versionen nur das Hochskalieren niedrig aufgelöster Bilder auf die höhere Zielauflösung ermöglichte. „DLSS 3.0“ ist deshalb jedoch limitiert auf die aktuell neueste Version von NVIDIA-Grafikkarten (Ada Lovelace) und demnach nicht kompatibel mit Grafikkarten anderer Hersteller.

AMD hingegen stellte mit „FSR“ eine offen zugängliche Alternative vor, welche keine besonderen Hardwareanforderungen besitzt. Dies liegt daran, dass „FSR“ im Gegensatz zu „DLSS“ kein neuronales Netz zur Optimierung verwendet, sondern einen Algorithmus,

der lokal auf den Shader-Einheiten einer GPU durchgeführt werden kann.

„DLSS“ ist hier auf spezielle Rechenkerne (Tensor-Kerne) angewiesen, um die trainierten Ergebnisse des neuronalen Netzes in Echtzeit anwenden zu können, was die Technologie auf NVIDIA-Grafikkarten limitiert (Andemahr, 2022).

Dadurch schaffte es „FSR“ laut AMD zu einer doppelt so schnellen Verbreitung wie „DLSS“ innerhalb der letzten zwei Jahre (Blake-Davies, 2022). Ähnlich wie „DLSS“ war im Jahr 2021 die erste Version von AMDs „FSR“ von ausbaufähiger Bildqualität – ab Version 2.0 konnte AMD die Lücke zu NVIDIA aber deutlich schließen.

Derzeit unterstützen ca. 168 Games AMDs „FSR“-Upscaling, während es über 200 Games sind, die NVIDIAS „DLSS“ unterstützen (AMD, 2023; NVIDIA, 2023).

Da Intels „XeSS“ eine vergleichsweise junge Image-Upscaling Methode ist, kommt die Anzahl der unterstützten Spiele derzeit auf 50 (Stand: März 2023, Intel, 2023).

6. Herausforderungen

Deep Learning Image Upsampling Technologien haben trotz der vielen Vorteile auch einige Herausforderungen. Zum einen können beim Upscaling von Bildern Artefakte wie Unschärfe, Kantenrauschen und Blockbildung auftreten. Diese Artefakte können die Bildqualität beeinträchtigen und das visuelle Erlebnis negativ beeinflussen.

Die Verwendung von Machine Learning-Techniken und Deep Learning-Techniken zur Verbesserung der Bildqualität erfordert außerdem eine große Menge an Trainingsdaten, welche für das neuronale Netz benötigt werden. Dies kann die Komplexität des Upscaling-Prozesses erhöhen und die Implementierung erschweren. Besonders die verwendeten Algorithmen, zur Verfügung stehende Supercomputer zur Verarbeitung der Trainingsdaten für neuronale Netze (im Falle von DLSS beispielsweise) wirken sich auf die Qualität der Image-Upscaling Techniken aus.

Beim Upscaling von Bildern können außerdem feine Details verloren gehen, insbesondere bei der Verwendung von einfachen Methoden wie der Bilinear Interpolation. Dies kann die wahrgenommene Bildqualität und Schärfe des Bildes beeinträchtigen.

7. Fazit und Ausblick

Zusammenfassend lässt sich sagen, dass Image-Upscaling mit Deep Learning das Potenzial hat, die Spieleindustrie nachhaltig zu verändern. Durch die Verbesserung der Bildqualität und die Verringerung der Hardwareanforderungen können Spieleentwickler ein breiteres Publikum ansprechen und den Spielern ein noch intensiveres Spielerlebnis bieten. Da sich die Deep-Learning-Algorithmen weiter verbessern, kann man in den kommenden Jahren mit noch größeren Fortschritten bei der Bildhochskalierung rechnen.

Darüber hinaus bietet diese Technologie das Potential der Einsparung von Rechenleistung, was auch auf ökologischer Ebene Signifikanz hat und den Stromverbrauch eines Computers reduzieren kann. Hierbei sei ein leichtes Gegensteuern der wachsenden Leistungsaufnahme insbesondere bei GPUs denkbar.

Eine Gegenüberstellung der verschiedenen Image-Upscaling Techniken macht vor allem Sinn, wenn man Bewegtbildinhalte betrachtet. Im Anhang befinden sich deshalb einige Gegenüberstellungen der aktuellen Image-Upscaling Methoden.

8. Literaturverzeichnis

- [1] AMD. (2023). *AMD FidelityFX™ Super Resolution*.
<https://www.amd.com/en/technologies/fidelityfx-super-resolution>
- [2] Andemahr, W. (2022, Mai 12). *AMD FSR 2.0, FSR 1.0 und Nvidia DLSS in Vergleich—ComputerBase*. <https://www.computerbase.de/2022-05/amd-fsr-2.0-deathloop-test/>
- [3] Blake-Davies, A. (2022, Juni 22). *AMD FidelityFX Super Resolution at One Year: 110+ Games and FSR 2.0 Now Available for Developers*. AMD.Com.
<https://community.amd.com/t5/gaming/amd-fidelityfx-super-resolution-at-one-year-110-games-and-fsr-2/ba-p/531875>
- [4] Brownlee, J. (2019, Juni 16). A Gentle Introduction to Generative Adversarial Networks (GANs). *MachineLearningMastery.Com*. <https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/>
- [5] Caulfield, B. (2018, März 19). *Ray Tracing | NVIDIA Developer*. <https://developer.nvidia.com/discover/ray-tracing>
- [6] chaitanya. (2021, Juni 19). *What is Anti-Aliasing? Different Types of Anti-Aliasing Techniques*. Electronics Hub. <https://www.electronicshub.org/anti-aliasing/>
- [7] Davenport, M., Flesher, C., Poonawala, M., & Suksumrit, D. (o. J.). *Motion Vector Applications—Home*. Motion Vector Applications. Abgerufen 3. März 2023, von <https://www.clear.rice.edu/elec301/Projects02/motionVector/home.html>

- [8] Deng, L. (2014). Deep Learning: Methods and Applications. *Foundations and Trends® in Signal Processing*, 7(3–4), 197–387.
<https://doi.org/10.1561/20000000039>
- [9] Laricchia, F. (2022, Februar 14). *4K Ultra HDTV household penetration in U.S. 2021*. Statista. <https://www.statista.com/statistics/1247334/4k-ultra-hdtv-us-household-penetration/>
- [10] Lu, P., Morris, M., Brazell, S., Comiskey, C., & Xiao, Y. (2018). Using generative adversarial networks to improve deep-learning fault interpretation networks. *The Leading Edge*, 37(8), 578–583. <https://doi.org/10.1190/tle37080578.1>
- [11] Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.
- [12] NVIDIA. (2023). *NVIDIA DLSS-Technologie*. NVIDIA.
<https://www.nvidia.com/de-de/geforce/technologies/dlss/>
- [13] Pratt, M. (2023, Februar 21). *What is 4K TV? Which?*
<https://www.which.co.uk/reviews/televisions/article/tv-screen-technology-explained/what-is-4k-tv-aA9be0m2H18T>
- [14] Salian, I. (2020, Februar 3). *What Is AI Upscaling?*
<https://blogs.nvidia.com/blog/2020/02/03/what-is-ai-upscaling/>
- [15] Schiesser, T. (2020, Oktober 2). *Ray Tracing & DLSS with the GeForce RTX 3080*. TechSpot. <https://www.techspot.com/article/2109-nvidia-rtx-3080-ray-tracing-dlss/>

- [16] Stanar, D. (2020, November 19). *Best Anti-Aliasing | Different Methods to Get Clearer Images*. <https://www.vssmonitoring.com/best-anti-aliasing/>
- [17] Syed, A. (2021, Juni 30). AMD FidelityFX Super Resolution vs Vanilla Upscaling + LumaSharpen: Is FSR Better than Traditional Spatial Upscaling Techniques? *Hardware Times*. <https://www.hardwaretimes.com/amd-fidelityfx-super-resolution-vs-vanilla-upscaling-lumasharpen-is-fsr-better-than-traditional-spatial-upscaling/>

9. Anhang

Gegenüberstellung verschiedener Image-Upscaling Methoden:

[1] <https://youtu.be/D4atsxcLQ-o>

[2] https://youtu.be/7Y_TnWybvU8